

Bridging the Taiwan Strait

The political and practical obstacles to greater scientific exchange between Taiwan and the Chinese mainland are significant. It is achievable, however, and is in the best interests of both parties.

The election earlier this year of Taiwan's new president Chen Shui-bian, whose Democratic Progressive Party contains radicals advocating independence, has brought political tensions with the Chinese mainland to new heights, almost verging on war. Vitriolic attacks by the mainland on pro-independence members of Chen's administration have created an atmosphere in which attempts by well-intentioned individuals, such as Nobel prizewinner Yuan-tseh Lee, to increase scientific exchange with the mainland are being seriously hampered (see *Nature* Regional Insight, pages 415–426). This benefits no one.

Taiwan has a well-developed infrastructure for research at its universities and government research institutes, and at Academia Sinica, Taiwan's premier institution for basic research headed by Lee. But these institutions are facing a manpower crunch. A flood of returnee scientists from the United States in the early 1990s has been reduced to a trickle as the US economy picks up and tensions with the Chinese mainland rise. Furthermore, students in Taiwan are turning their backs on PhD and postdoctoral research, preferring to take up positions in Taiwan's booming electronics industry or to study business management or economics rather than science.

This offers a golden opportunity for exchange with the mainland if political differences could be put aside. The Chinese mainland has a large supply of young high-quality scientists willing to do basic research. And leaders of Taiwanese research institutions have growing respect for the quality of scientists on the mainland, who still have to work under comparatively poor conditions.

Nor need the exchange be confined to basic research. One entrepreneurial scientist at Tsinghua University in Beijing, for example, has set up a start-up company for biochips in California funded with venture capital from Taiwan and plans to manufacture the chips on the island (see page 426).

At present, there are about 50 mainland Chinese scientists on short-term visits at Academia Sinica. Lee, who maintains strong

links with mainland institutions and is much respected there, actively encourages mainlanders to come to Taiwan. But the duration of their stay is limited to a maximum of one year by mainland regulations. Furthermore, the visiting scientists are not issued with passports because the mainland government considers Taiwan a renegade province. This prevents the visitors from, for example, attending conferences outside Taiwan, and severely limits the value of their stay. Allowing longer stays of at least two years would be a simple step that would benefit both parties.

But it would be naive to suggest that opening the taps on exchange alone will do much to improve scientific relations between Taiwan and the mainland. Some fundamental problems still stand in the way of joint research. Many institutions in Taiwan, such as Academia Sinica and National Tsinghua University, have adopted the names of the mainland organizations from which they claim their roots. Furthermore, the word 'National' tagged onto many research institutes and university names in Taiwan, such as National Taiwan University, is like a red rag to a bull for officials on the mainland, who are vehemently opposed to any suggestion that Taiwan is a nation. This presents very practical problems if Taiwan's researchers and their colleagues on the mainland want to publish joint research in international journals. The names in the affiliations of Taiwan's researchers are often simply unacceptable to mainland organizations and the government that funds them — and in some cases, such as that of Tsinghua University, the names are the same.

Yet, with goodwill and imagination on the part of both parties, solutions to this sort of problem can be found — as, for example, can be seen in Taiwan's participation in the current Sydney Olympics under the name of Chinese Taipei.

It is essential to improve and increase scientific exchanges, and not just because it would benefit science in both Taiwan and the Chinese mainland. Anything that helps ease the political tensions across the Taiwan Strait is good for the world as a whole. ■

Reviews for reviews' sake

Next month sees three new *Nature* review titles that make the most of the partnership between editors and authors.

Journals that aim to take an overview of research have a long and honourable history. The world's longest-running journal, the *Philosophical Transactions of the Royal Society*, was at its outset in 1665 arguably a review journal of sorts, aimed as it was at "giving some ACCOUNT of the present undertakings, studies, labours of the **INGENIOUS** in many considerable parts of the **WORLD**".

Nature only began to publish review articles with any frequency in the 1970s, but now *Nature* and all of the monthly *Nature* journals see them as an essential part of their content. Feedback suggests that readers welcome *Nature*'s most recent development in this direction: the *Nature* Insight collections that appear about once a month.

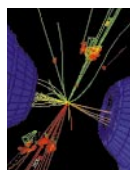
But *Nature*'s publishers have taken a significant step further. On 1 October, visitors to <http://www.nature.com/reviews/> will find three new journals, under the banner of *Nature Reviews*, devoted to

overviews of recent highlights in genetics, neuroscience and molecular cell biology. Information on these journals can be found elsewhere in this issue.

What should perhaps be emphasized above all is the combination of quality of authors and the depth of attention given to the review articles by their editors, in terms of topics and content, readability, addition of complementary material on the web, and in the figures and web-based animation. The journals' philosophy is fully in keeping with the tradition of *Nature* journals, by which much of the creativity starts — and the buck stops — with the editors themselves, while the authors deliver the clout.

With many questions being asked about the addition of value by publishers, we intend for these partnerships between editors and authors of reviews to represent outstandingly positive examples. ■





Holy Grail?
Higgs boson 'events' win LEP collider a temporary reprieve
p276



Mercury rising
European Space Agency urged to send mission to planet
p277



Scrutable
Mouse cDNA marks new openness in Japanese genomics
p279



On a NIF edge
Budget estimate throws fusion project into doubt
p280

Doubts grow over discovery of fossilized 'dinosaur heart'

Rex Dalton, San Diego

This spring, a team led by researchers from North Carolina State University received widespread publicity for their claim to have discovered the remains of a heart within the ribcage of a 66-million-year-old dinosaur specimen.

But since then doubts have been growing about whether the discovery is all that it seemed. Palaeontologists from the University of Texas are now preparing a rebuttal that is expected to say that the fossil — from a herbivorous *Thescelosaurus* — is a deposition of minerals known to geologists as a concretion.

Nature has also learned that one of the original article's authors — a commercial fossil hunter who found the specimen in South Dakota — has, according to federal records, been convicted of trafficking in Native American artefacts stolen from federal lands.

Although there is no evidence of any wrong-doing in the current case, and no one is making any such claims, the fact that the researchers were unaware of this background has again raised the question of

how thoroughly institutions investigate the provenance of the material they collect.

The heart was described as having four chambers and a single systemic aorta, more like a human heart than that of a reptile, which has only three chambers and paired systemic aortas. This suggested to the authors "the existence of intermediate-to-high metabolic rates among dinosaurs", strengthening the hypothesis that many were warm-blooded (see *Science* **288**, 503–505; 2000).

But immediately after the article was published, some top US palaeontologists questioned its claims. The team at North Carolina State University — including lead author Paul Fisher and senior curator of palaeontology Dale Russell — then put up computerized tomography (CT) scan photos of the specimen on the museum's website.

Last month, says Russell, the specimen was inspected by palaeontologist Timothy Rowe, an authority on CT scans and director of the Vertebrate Paleontology Laboratory at the University of Texas, the only facility with a CT scanner dedicated to such studies.

Working with University of Texas geolo-



Heartless? The fossil dinosaur's 'heart' (centre) may actually be a mineral concretion.

gist Earle McBride, an authority on palaeontological concretions, Rowe is preparing the rebuttal for *Science*. Rowe — reportedly the only external palaeontologist to see the specimen since it was described — declines to discuss the details of his report before submission and publication.

But Russell says that, after visiting the museum, Rowe "doesn't think it is a heart". "We have no fear," adds Russell. "We will follow the truth. We have no case but to find out what it is. I look at it as an adventure."

North Carolina officials say they were unaware of the court record of Michael Hammer, the fossil hunter who provided the specimen. Hammer, who runs Hammer and Hammer Paleotek, a palaeontology preparation company near Jacksonville, Oregon, has refused several requests for an interview.

Officials from the North Carolina Museum of Natural Sciences in Raleigh — where the heart specimen is displayed in a \$70 million facility opened last April — maintain that their enquiries when they bought the fossil in 1996 were sufficient to document that the sale was legitimate.

Russell says he had known Hammer for more than five years before that, and that

UK advised to step up asteroid hunt

Natasha Loder, London

The British government may take a bigger part in international efforts to identify objects in space that threaten the Earth. Such a move has been urged by an independent panel known as the Near Earth Objects Taskforce set up by the science minister, Lord Sainsbury, whose report was published in London this week.

Other recommendations include building a three-metre telescope in the Southern Hemisphere, and dedicating the one-metre Johannes Kapteyn Telescope on La Palma in the Canary Islands to following up observations of near-Earth objects (NEOs).

Sainsbury said that although the possibility of a dangerous NEO hitting the Earth was "extremely remote", it had potentially serious consequences. "We put a

lot of money into astronomy, and I think it is sensible to put a bit into making certain that we know if there is any danger of an object hitting our very fragile planet."

He added that he plans to consult with ministerial colleagues and international bodies about how the government should respond to the committee's proposals.

The taskforce argues that international missions could detect NEOs while making other observations. It says the government should ask the European Space Agency to ensure that one of its future missions also surveys the sky for such objects.

The taskforce also says that the Minor Planet Center at the Smithsonian Astrophysical Observatory in Cambridge, Massachusetts, should be put on a "robust international footing".

▶ Hammer invited him to the fossil and mineral show in Tucson, Arizona, in about 1995 to see the *Thescelosaurus* heart, which he was trying to sell. Hammer has said he found the fossil in 1993 in the Hell Creek formation on private land in northwest South Dakota.

Using funds raised by a support group, the North Carolina museum paid Hammer \$350,000 for the *Thescelosaurus*. When told that federal records showed that Hammer pleaded guilty in September 1994 to two charges of trafficking in artefacts stolen from federal land in Utah, Russell said he was unaware of this.

In 1990, when federal and state authorities raided Hammer's home, records show they found Native American human remains, which were repatriated to a Californian tribe in 1996. Through his wife, Hammer has denied any impropriety over the *Thescelosaurus* specimen. She attributes the earlier prosecutions to overly aggressive law enforcement in an area of changing, confusing legislation.

Earlier this month, North Carolina palaeontologists took CT scans of the heart specimen at the university's veterinarian college. Museum officials say they are clearer and more helpful than the original scans. But some palaeontology authorities say no conclusion can be drawn until the remnant is cut open. ■

▶ <http://www.dinoheart.org>

World Bank is urged to give greater priority to science

Colin Macilwain, Washington

A prominent US economist has criticized Western governments and the World Bank for failing to acknowledge the full potential of science, technology and innovation in alleviating global poverty.

The criticism has come from Harvard University professor Jeffrey Sachs, director of the Center for International Development and former director of the Harvard Institute for International Development. He told the President's Committee of Advisors on Science and Technology (PCAST) last week that the World Bank's policies, based on the premise that open markets and good government would create growth in poor countries, were "vastly inadequate".

"There's a huge missing piece" in conventional thinking about development, Sachs told the PCAST meeting. "That is the recognition that technological change is the main driver of economic growth, but that it doesn't emerge endogenously in tropical countries."

The attack was on the *World Development Report 2000/2001* — released by the World Bank on 12 September — which emphasizes "opportunity, empowerment and security" as the keys to poverty alleviation. "There's no recognition in this document of where the



Sachs: critical of policies to tackle global poverty.

real problems lie," he claimed.

Sachs proposed three ways to mobilize science and technology in developing countries: the direct transfer of technologies, such as vaccines, from rich countries; the spread of

technology through the globalization of manufacturing; and the construction of science and technology development capacity in the developing countries themselves.

World Bank officials responded vigorously to the charges. Nora Lustig, director of the *World Development Report* at the World Bank, said in an interview that the importance of science and technology was repeatedly acknowledged in the document.

The World Bank is currently developing a new science and technology strategy. And supporters of science at the bank say they are gradually winning support for the idea that science and technology are vital to development, even in the poorest countries. ■

▶ <http://www.worldbank.org/poverty/wdrpoverty>

Hunt for Higgs particle wins time for CERN collider

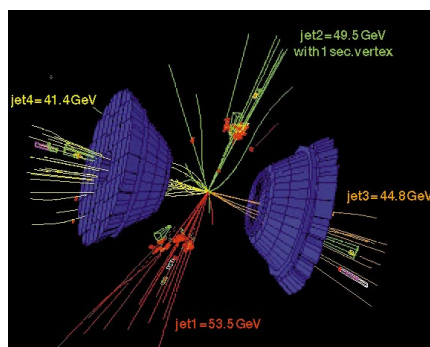
Alison Abbott

Physicists at the European Laboratory for Particle Physics (CERN) near Geneva believe that a new analysis of their recent observations indicates an increased probability that they have witnessed the transitory creation of their current Holy Grail — the Higgs boson.

Although they do not rule out the possibility that there may be another explanation for the events they have observed, the analysis helped physicists last week to convince CERN's director-general, Luciano Maiani, to extend the operation of its Large Electron-Positron (LEP) collider by a month.

LEP was to have been switched off at the end of this month to allow construction to start on its successor, the Large Hadron Collider (LHC). Keeping LEP running for an extra month will cost Sfr7 million (US\$4 million). This will be found by shuffling CERN's existing budget, although Maiani says it will not disrupt LHC's schedule, or increase its cost.

Scientists working on LEP's four detector



Tell-tale signature? These four jets of quarks may come from the decay of a Higgs boson.

experiments had requested a temporary reprieve because their pooled results had pinpointed five 'events' that could be interpreted as evidence for the existence of the elusive particle. The Higgs boson is thought to be produced when electrons and positrons collide at an energy of around 115 giga-electron volts (see *Nature* 407, 118; 2000).

Re-analysis of the results shortly before last week's decision had produced a fivefold

reduction in the odds — to 1 in 1,000 — that the observed decay patterns, or 'signatures', are random events. But these odds are still four orders of magnitude too high for physicists to say that the existence of Higgs has been demonstrated conclusively.

LEP's temporary reprieve will allow its scientists to double the amount of data they can collect from particle collisions at the high energy levels that LEP was originally designed to achieve. If the signatures of the boson continue to be seen, the confidence limits will have been increased by two orders of magnitude.

"We are excited," says Tiziano Camporesi, head of LEP's DELPHI experiment. But he points out that the observations still have to be confirmed not only with a higher degree of confidence, but also in a way that is consistent across all detector experiments and all decay modes of the particles involved in the collisions. So far, no events have been seen in other decay modes apart from those involving the creation of quarks. ■

▶ <http://cern.web.cern.ch/CERN>

European space panel picks new priorities

Alison Abbott, Munich

The European Space Agency (ESA) is being asked to approve a package of missions that includes a major project to Mercury and a significant part in the NASA-led Next Generation Space Telescope (NGST). The move is intended to meet the needs of space scientists across a range of disciplines.

Both projects received the backing of ESA's Space Science Advisory Committee (SSAC) last week, and are scheduled for launch in about 2009. The panel also approved a solar mission, the Solar Orbiter, and an astrometry mission, GAIA, both scheduled for launch around 2012.

If approved — as is widely expected — the mission to Mercury will be ESA's fifth major 'cornerstone' mission, with a budget of around 600 million euros (\$514 million). It is known as Bepi Colombo, after the Italian space pioneer Giuseppe Colombo who in the 1960s worked out how to use the gravity fields of planets to 'slingshot' a spacecraft towards its destination, a technique used on NASA missions to Mercury in the 1970s.

Bepi Colombo will use a new solar-electric propulsion system. It will have a lander to investigate Mercury's surface and remote-sensing instruments to study the planet's surface and magnetosphere.

ESA will contribute around 150 million euros to the NGST — about 15% of its estimated total cost. This is comparable to ESA's involvement in the Hubble Space Telescope.

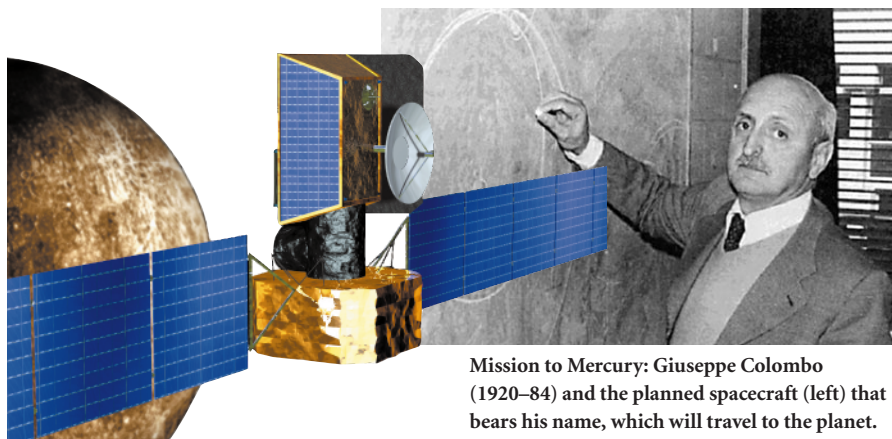
Bernard Seery, NGST's project manager at the the Goddard Space Flight Centre near Washington DC, says he is delighted but not surprised by the recommendation. "ESA and NASA have been working together on the development of the NGST concept since its inception," he says.

The Canadian Space Agency will take a 5% stake in the NGST, and some European countries may increase their participation. Italy has offered to add a further ten million euros, to be handled by ESA, and the United Kingdom is believed to be negotiating a bilateral agreement with NASA.

As with Hubble, ESA will contribute to the spacecraft as well as its instruments. It is slated to provide the NGST's near-infrared multi-object spectrometer, an innovative apparatus that will allow about 100 celestial objects to be observed at once.

In juggling the interests of researchers working on the Solar System, astronomy and fundamental physics, the SSAC decided to recommend that GAIA — another cornerstone mission — be launched unusually soon after Bepi Colombo.

ESA's earlier astrometry mission, Hipparcos, completed its data collection in 1997, and ESA is keen to ensure that the groups analysing these data do not disperse before



Mission to Mercury: Giuseppe Colombo (1920–84) and the planned spacecraft (left) that bears his name, which will travel to the planet.

GAIA is developed. "A delay in its development would be a risk for the continuity of these groups," says Jordi Torra, an astronomer at the University of Barcelona.

The SSAC recommended the Solar Orbiter as its third 'flexible' mission, rather than a Mars mission as some had predicted. This was partly out of concern that taking on Bepi Colombo on top of a number of other large missions — including Mars Express and the cornerstone mission to comet Rosetta, both due for launch in 2003 — might overstretch Europe's planetary scientists.

Also, the joint NASA/ESA mission to Sat-

urn, Cassini-Huygens, launched in 1997, will reach its destination in 2004. "Research groups must be available to take on these projects," says Giovanni Bignami, science director of the Italian Space Agency. "There is some concern that in planetary sciences they are getting saturated."

The SSAC also recommended that ESA uses a future 'flexible' mission to participate in another planned NASA-led project to detect gravity waves. ESA's Space Policy Committee will decide formally on the new missions next month.

Additional reporting by Xavier Bosch, Barcelona.

NASA reaches out to universities

William Triplett, Washington

The US space agency NASA is trying to bring more university researchers — with fresh ideas — into its fold. But although they welcome the initiative, many researchers are waiting to see how much money NASA will put where its mouth is.

In recent years, budget cuts have forced NASA to reduce its research spending. As a result, universities have felt neglected, and have criticized the agency for becoming too inward-looking. The new effort, dubbed the 'University Initiative', is intended to change that, says Randall Correll, a NASA spokesman.

Spence Armstrong, a US Air Force veteran who has been with NASA since 1991 — most recently as head of the agency's aeronautics and space transportation section — will oversee the initiative. In February he was appointed senior adviser to NASA administrator Dan Goldin.

The scheme will be launched next month, when NASA will hold the first of several planned 'cyber-conferences' in Washington. The conference, to be broadcast live over a NASA cable television

channel as well as the Internet, is intended to develop ideas for collaboration in four areas: space flight, space science, Earth science and aerospace technology.

During the conference, Armstrong and Goldin will invite proposals from university researchers. Although they must ultimately benefit NASA, proposals will not have to be strictly space-related.

"If you look at recent NASA activities, some of the technologies we're starting to invest in, such as biologically inspired engineering, these types of things show that we're expanding," says Correll. "Things have become much more complicated than just building a rocket ship and going into space."

For example, says Correll, a robotic mission to Mars to search for signs of life might make use of studies of life on Earth. "Once you've studied what life is like here, you might change your type of sensor sweeps that you put on your spacecraft."

But despite a 7% increase in NASA's budget, Armstrong acknowledges that the agency will need even more money in the next budget to underwrite the initiative. He will not name a figure, but he has told his



Armstrong: looking for a bold budget request.

budget designers to be "reasonably bold" in their request.

Another variable is the US presidential election in November. President Bill Clinton's administration will prepare the next federal budget before

then, but Congress will not take it up until after the new administration proposes any adjustments.

But some observers are optimistic. "Armstrong has already come to us and asked our membership for ideas and advice," says Jacques L'Heureux, director of University Relations for Space Sciences at the Universities Space Research Association. "He seems very serious."

University researchers have been "very receptive", says Correll. But, he adds, "the biggest fear, as you might expect, is that when the government says 'we're here to help', the feeling is that talk is easy."

♦ <http://www.hq.nasa.gov/office/codea/codea/WWW/news.html>

Germline gene therapy needs tight control, says US panel

Paul Smaglik, Washington

Tight regulation and oversight will be needed of efforts to make genetic modifications to human cells that can be passed on to offspring, a panel set up by the American Association for the Advancement of Sciences said this week.

Proponents of germline gene therapy have previously asked that it should not be regulated more heavily than other kinds of experimental medicine (see *Nature* 392, 317; 1998). Such a move, they argue, could slow efforts to develop germline gene therapy, which aims to combat genetic diseases by targeting cells in a prospective parents' testes or ovaries, or by conducting gene therapy on an embryo.

But the recommendation of the AAAS panel reflects its conclusion that, for all its promises, germline therapy remains problematic because unintended genetic changes could be passed to a new child along with intended benefits.

"There would need to be compelling scientific evidence to prove that these procedures are safe and effective," says Mark

Frankel, the report's co-author and director of the AAAS Scientific Freedom, Responsibility and Law programme. Such evidence has not yet materialized.

Rather than pursuing germline therapy, the AAAS report urges scientists to focus on making changes in cells that will not be passed to the next generation. It also encourages more research into methods to correct genetic 'misspellings'.

Without first improving either of those two approaches, germline gene therapy can be neither safe nor effective, Audrey Chapman, director of an AAAS program on biomedical ethics and a co-author of the report, said at a press conference on Monday.

Robert Cook-Deegan, a Georgetown University bioethicist, called for a new body to be set up to monitor germ line gene therapy experiments in animals, perhaps modelled on the National Institute of Health's Recombinant DNA Advisory Committee (RAC). It would eventually decide whether human clinical trials should go forward.

♦ <http://www.aaas.org/spp/dspp/sfrl/germline/main.htm>

Case of the stolen Enigma machine takes cryptic turn

Natasha Loder, London

It is a riddle wrapped in a mystery inside an Enigma. Who stole a rare version of the machine that the Germans used to encrypt military secrets during the Second World War? And does an anonymous note contain a coded clue to the thief's identity?

The 'Enigma machine' was stolen in April from Bletchley Park Museum in Buckinghamshire. During the war, Bletchley Park was home to a team of mathematicians and cryptographers who worked on cracking coded German military communications.

In the past weeks the museum has received two strangely worded, anonymous letters offering to return the machine for a five-figure sum. The first starts: "I have been asked by the current owner the above Enigma machine, who purchased it in good faith (in good faith being the operative words) to say and tell you now today, the unwitting person have no desire of depraving [sic] your august self or any one the pleasure to see it again." It was sent from the Midlands of England, and appears to have been written on a typewriter dating from the Second World War.

The museum's chief executive, Christine Large, says the episode has "an element of mystique which is completely in keeping



with the history of the park. It is an unusual letter but seems genuine. It doesn't surprise me that the author is a middleman. We really want to assure them that our priority remains getting [the machine] back in good condition."

The museum is now waiting for the next contact. "We have made it very clear that, assuming there is no criminality involved, that police will not prosecute," says Large. The machine was on long-term loan from the Government Communications Head-

quarters in Cheltenham, which acquired it after the war.

There are several hundred Enigma machines on the market. But the stolen one, which carries the serial code G 312, is a very rare four-rotor machine used by the Abwehr German military intelligence. The only other known example is owned by the US National Security Agency.

Bletchley Park became the home to the UK Government Code and Cypher School in 1939; the teams that worked to crack the Enigma machines were housed in wooden huts. The Germans updated the machines over the years, but a series of breakthroughs by cryptographers meant that German messages could be decoded for much of the war.

The Buckinghamshire police say the letters are their best lead so far. The first letter ends in a word — which is not being released by police — that some have suggested may be in code. But Large says that the word is recognizable, although not English. The machine featured in an Interpol 'wanted' poster in June, but the police say they have no evidence that it has left the country.

♦ <http://www.bletchleypark.org.uk>

♦ <http://www.interpol.int/public/workofart/data/1003/1003713.asp>

Japan opens access to mouse cDNA data...

David Cyranoski, Tokyo

Japanese research institutes are not known for their openness towards foreign collaborators. But a growing movement to change this was signalled by a two-week meeting in Tsukuba earlier this month, held to annotate a large set of mouse complementary DNA held by the Institute of Physical and Chemical Research (RIKEN).

Unusually for a Japanese research institute, RIKEN's Genomic Sciences Center has initiated and taken the lead in a collaboration based on the free public distribution of data and research materials.

Yoshihide Hayashizaki, a chief scientist at the centre and leader of the project, invited about 50 bioinformaticists and biologists to the Functional Annotation of Mouse (FANTOM) meeting to annotate a subset of his library of 128,500 mouse complementary DNAs, which he says is the world's largest set of mammalian cDNAs.

Many foreign participants say they were struck not only by the event's scale, but also by its taking place in Japan at all. "Japan has always been so secretive," says one US researcher. Another attributed the internationalism to Hayashizaki's efforts.

Hayashizaki says he faced a catch-22 in getting government approval for the project. "We are told to get involved in international collaborations and to make a name for Japanese science on the international scene, yet we are expected to keep results paid for by Japan-

ese taxpayers, and criticized if we just give them away for other countries to profit from."

Unlike an entire genome, in which genes are interspersed with 'junk' DNA, cDNA is made from RNA produced in the cell, so each sequence represents an expressed gene.

Annotation of cDNA sequences aims to predict the structure, function, homology and protein-protein interactions of their gene products.

Huddled around computers at some times and karaoke machines at others, the group annotated about 21,000 full-length cDNAs. After removing redundant or previously known sequences, this left just under 16,000 unique cDNAs.

Researchers say they have already discovered a number of new genes, or genes that were previously unknown in mice or mammals. The similarity between the mouse and human genomes means that the mouse will almost certainly provide important clues to understanding the human genome.

The data will no doubt find profitable application. Full-length cDNA are "notoriously hard to get," says participant Peter Mombaerts, a neuroscientist at the Rockefeller University in New York, as the RNA can easily degrade during collection or form secondary structures that make it difficult to copy into cDNA.

After the results are published, Hayashizaki plans to release the data to two public databases, the DNA DataBase of Japan and



Embracing change: Hayashizaki, left, and his RIKEN colleague, Yasushi Okazaki.

the Mouse Genomic Database at the Jackson Laboratory in Bar Harbor, Maine. The cDNA clones will be made available to researchers worldwide.

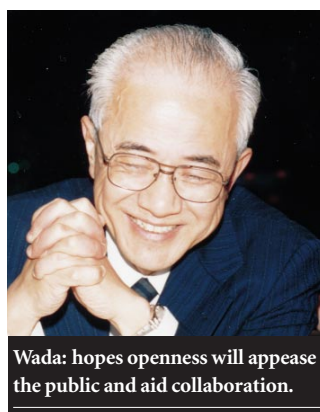
Hayashizaki believes that his clones will become the standard for future research, and he is already planning FANTOM II. Many emerged from this month's meeting convinced that he has the best mouse cDNA around, and that the meeting marked a significant step for Japanese science in keeping pace with collaborative trends elsewhere. ■

...and to report on strengths and weaknesses of genomics centre

Anyone who wants to know what is wrong — and right — with research at the Genomic Sciences Center (GSC) of the Institute of Physical and Chemical Research (RIKEN) can turn to RIKEN's website. There they will find a detailed report containing both praise and criticism of the centre.

The material reflects the keenness of the GSC's director, Akiyoshi Wada, both to meet the Japanese public's demand for transparency in government activities and to create the openness that he feels is "necessary for interaction with the international scientific community".

In March an external advisory committee, made up of four foreign and five Japanese researchers, reviewed the GSC's five research groups. Wada, who has been GSC director for two years, has made the unconventional decision to post its



Wada: hopes openness will appease the public and aid collaboration.

report, along with his item-by-item response, on RIKEN's website.

Panel member Kurt Wüthrich of the Institute for Molecular Biology and Biophysics at ETH-Hönggerberg in Zürich says Wada has made "a bold move". "I have served on several such committees in France and Germany, and

they always make the findings 'public', but that usually just means to make public within the research institute."

The overwhelmingly positive comments — including "number one in the world" for Yoshihide Hayashizaki's project on mouse complementary DNA (see above) and "world leader, cannot be criticized" for Yoshiyuki Sakaki's Human Genome Research Group — perhaps made the decision to go public an easy one.

But Wada also had to address some delicate issues. For example, the committee suggested ways to make the GSC's new nuclear magnetic resonance facility, used for analysing protein structure, an "internationally visible, open facility".

Wada claims that it is impossible to meet these suggestions, but he is working out a compromise in which about 10% of the facility's time will go

to researchers in Japan and abroad who have been selected by international committee.

Some of the research-group directors are worried that making the evaluations public could perpetuate misunderstandings. For example, Hayashizaki praised the efficiency of his "factory-like" cDNA sequencing efforts. But the report turned this description on its head, with an implication that his work is factory-like and thus not scientific.

Some government officials are worried that its international outlook prevents GSC from repaying taxpayers with 'economic fruit'. Wada replies: "There are two ways to make fruit: we could try to make money or we could raise the level of Japan's scientific reputation. The second is much greater than monetary profit."

D. C.

♦ <http://www.riken.go.jp>

US Senate could pull funding plug on laser project

Washington The future of the National Ignition Facility laser project at the Lawrence Livermore National Laboratory in California is hanging in the balance this week after the Department of Energy sent Congress a final estimate saying that the project could be completed by 2008 at a total cost of \$3.5 billion.

This estimate assumes that Congress will add an extra \$135 million to the \$74 million budgeted for the facility's construction during 2001. But the Senate is losing patience with the project's cost overruns: on 7 September it passed an amendment capping such expenditure at \$74 million and ordering a study from the National Academy of Sciences into the project's relevance to the nuclear weapons programme.

The project's backers will now push for the extra money when House and Senate appropriators confer on their final budget for the Department of Energy. If the money is not forthcoming, "the project is in serious jeopardy," says a senior department official.

African trials begin of DNA vaccine for malaria

London Clinical trials of DNA vaccines for malaria, developed by scientists at the University of Oxford's department of medicine, began in the Gambia this week. Up to now, malaria vaccines have failed to produce a strong enough immune response. Instead of using inactivated malaria parasites or protein, a DNA vaccine uses fragments of the parasite's DNA.

Trials in mice have shown that a combination of the DNA vaccine and a vaccine containing a modified vaccinia virus Ankara produced a strong cellular — rather than antibody — response. High levels of T cells destroy infected liver cells. Phase 1 trials have now begun, with human volunteers testing the vaccine combination's safety and immune response.

Party urges Mbeki to toe the line on AIDS

Cape Town South African president Thabo Mbeki is coming under pressure from within his own political party, the African National Congress (ANC), to acknowledge publicly that HIV causes AIDS. The demand is made in a paper from the ANC's health committee that was leaked to the press, and comes shortly before Mbeki is due to receive a report on the proceedings of a panel set up earlier this year to look at the link between HIV and AIDS (see *Nature* 405, 105–106; 2000).

"The government's questioning of the link between HIV and AIDS is crippling the

campaign to combat the scourge," the weekly newspaper *The Mail and Guardian* wrote in an editorial last week. It further warned that either he "gets his act together on HIV/AIDS very soon or he places his presidency at risk".

Congress set to give NSF substantial budget hike

Washington The National Science Foundation (NSF) is heading for a substantial budget increase — although perhaps not as substantial as the record 17% hike proposed by President Bill Clinton's administration back in February — after the Senate approved a 10% increase last week that would take the agency's 2001 budget to \$4.3 billion.

Before a budget bill is sent to Clinton, the Senate must reconcile its proposal with the 4% increase approved by the House of Representatives for the agency in July. However, NSF supporters in both houses have said that more money should go to the agency if, as is likely, extra money becomes available from the appropriations bill, of which its budget is part, during late budget negotiations with the president.

Israeli fund will aid women researchers

Jersusalem The Hebrew University of Jerusalem has boosted prospects for women scientists with a US\$2 million fellowship fund that will give promising PhD candidates in the natural sciences US\$20,000 a year for three years. The fund has been set up by Dan Maydan, president of Applied Materials of Santa Clara, California, in memory of his wife, Dalia, a physical chemist who trained at the Hebrew University.

Maydan believes it is vital to encourage more women to go into science and technology, especially given the shortage of trained professionals in those fields in Israel. "I know very well the struggle women have to raise a family and pursue a scientific career," he said. Of the three recipients of the fund's first round of fellowships last week, one is a mother and another is pregnant.

Researcher in discrimination case settles with university

Montreal Kin-Yip Chun, the former research associate who has fought the University of Toronto for more than six years over allegations of racial discrimination, has reached an agreement with the university enabling him to return to seismological research (see *Nature* 397, 551; 1999).

The agreement follows a decision by the Ontario Human Rights Commission in July not to refer Chun's complaint to a board of inquiry. Chun will receive the title of research scientist and non-tenure track associate professor, start-up research funds, and the

right to teach and supervise graduate students and to compete for tenure-track positions.

Anthropology journal slashes prices

Paris Wiley-Liss, the publishers of the *American Journal of Physical Anthropology*, has agreed to cut the journal's annual subscription rate from US\$2,085 to US\$1,390 after negotiations with the American Association of Physical Anthropologists (AAPA). A second institutional subscription will be offered at 20% of the full rate, and the AAPA editorial office will also get more support.

"There's a very dramatic change going on in academic publishing," says Jonathan Friedlaender, professor of anthropology at Temple University, Philadelphia, and chairman of the AAPA's publications committee. "The era of very high charges is going to end either with cuts in prices like ours, or with many new competing journals owned by associations themselves."

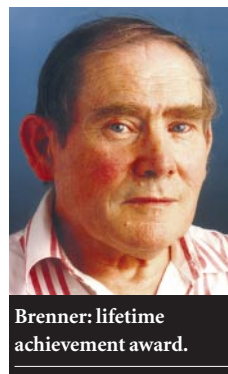
Proteins, virus and Brenner scoop Lasker awards

Washington Aaron Ciechanover and Avram Hershko, of Technion in Haifa, Israel, and Alexander Varshavsky of the California Institute of Technology, have won this year's Lasker prize for basic medical research for their discovery of the ubiquitin system of protein degradation. The prizes will be awarded in New York tomorrow (22 September).

The award for clinical medical research goes to Harvey Alter, of the National Institutes of Health in Bethesda, Maryland, and Michael Houghton, of the Chiron Corporation in Emeryville, California, for their work leading to the discovery of the virus that causes hepatitis C and the development of blood-screening methods

that have virtually eliminated the virus from the US blood supply.

The Lasker awards are often a pointer to the Nobel prizes. This year, Sydney Brenner, of the Molecular Sciences Institute in Berkeley, California — who



Brenner: lifetime achievement award.

has yet to win a Nobel despite a widespread belief that he should have done so — will receive the Lasker Award for Special Achievement in Medical Science. The Lasker Foundation acknowledges his "brilliant creativity in biomedical sciences" as well as his "rational voice" in the debate on recombinant DNA and his "trenchant wit".

A chemistry set for life

Small molecules that selectively disrupt the proteins encoded by individual genes could become powerful tools in functional genomics. Trisha Gura explores the nascent but highly promising field of chemical genetics.

Now would seem to be a good time to embark on a career in genetics. Demand will be high as laboratories the world over get to grips with the huge task of assigning functions to the thousands of genes in the human genome — and those of a growing number of model organisms. But the era of functional genomics could also create a rash of employment opportunities in chemistry.

Geneticists are realizing that chemists can make an important contribution by designing specific small molecules that can selectively disable particular genes and their protein products. Indeed, by combining genetics and chemistry, several research teams have managed to devise specific and reversible 'switches' to investigate the role of genes that have, until now, defied functional analysis.

This week's *Nature* includes a particularly elegant demonstration of this 'chemical genetic' approach¹. By introducing mutations into genes that encode for members of a class of enzymes known as kinases, a team led by chemist Kevan Shokat of the University of California at San Francisco can switch off individual kinases using specially designed chemical inhibitors. This is an important advance: kinase enzymes are involved in countless cell-signalling pathways, but biologists have lacked the tools to disable individual members of the family and so discern their precise functions. "Everybody who studies kinases will start using this," predicts Doug Kellogg, a cell biologist at the University of California, Santa Cruz, who is now collaborating with Shokat's team.

With other groups making good progress in applying similar approaches to different genes, chemical genetics should soon be yielding a rich harvest. "Our goal is to dissect the function of organisms and cells by having a small molecule partner for every gene product," says Stuart Schreiber of the Institute of Chemistry and Cell Biology (ICCB) at Harvard University. In addition to advancing our knowledge of fundamental biology, this approach could spur drug development — in many cases, chemicals that inhibit par-

ticular proteins could be developed into candidate drugs.

Not surprisingly, the US National Institutes of Health is starting to take an interest. Over the past two years, its largest centre, the National Cancer Institute (NCI) in Bethesda, Maryland, has awarded more than \$6 million to the ICCB and five other institutes working at the interface between chemistry and biology. "The discovery, modification and annotation of small molecules in terms of their ability to probe and perturb biological targets will be one of the central tools for the post-genomic era," says Richard Klausner, director of the NCI.

Family values

Shokat and his colleagues were drawn to chemical genetics through sheer necessity. Kinases participate in cell-signalling pathways by adding phosphate groups to other molecules. For their activity, they require adenosine triphosphate (ATP), a biochemical that acts as the energy 'currency' of the cell. But individual kinases guard the secrets of their function closely. If one kinase gene is mutated or deleted, disabling its enzyme product, other kinases often compensate over the course of the organism's development to assume the function of the disabled

enzyme. And although several chemicals can inhibit kinases, they are not specific to individual members of the enzyme class — you can disable the kinase you are interested in, but you will also knock out a handful of other kinases at the same time.

What Shokat needed was a way to disable individual kinases without giving cells enough time to call up a stand-in. The researchers began by mutating a selected kinase gene to change the site on the enzyme that binds to ATP. This enlarged the ATP-binding pocket but otherwise did not affect the enzyme's normal function. Then came the chemistry. The researchers added a variety of bulky chemical groups to small molecules already known to inhibit kinases, searching for a molecule that could slot neatly into the enlarged site and stop ATP from activating the enzyme.

Through this marriage of chemistry and genetics, Shokat and

Budding out: chemical genetics has shed light on an enzyme crucial to yeast reproduction.

DAVID SCHARF/SPL

his colleagues turned a non-specific kinase inhibitor into a highly specific one. "We combined genetics and chemistry in a new way that allows you to study kinases with a very regulatable switch," says Shokat. And the switch is reversible, unlike a permanent mutation. Once the inhibitor is removed from the fluid bathing the cells, it is gradually replaced by ATP in the affected kinase, which turns the enzyme back on.

The team reported their first successes in 1998, disabling two kinases². They have since expanded on this work^{3,4}, and their *Nature* paper provides evidence that the approach can be applied to any kinase¹. Using chemical derivatives of two different kinase inhibitors, Shokat's team has, to date, disabled kinases from five different sub-families, working with the yeast *Saccharomyces cerevisiae*, mammalian cells and mouse embryos.

Already, Shokat and his colleagues are uncovering information on the function of kinases. With Dave Morgan, a cell biologist at the University of California at San Francisco, they have gained fresh insights into the function of Cdc28, the major kinase regulating the cell cycle in yeast. And in the latest issue of *Nature Cell Biology*⁵, working in collaboration with David Drubin of the University of California at Berkeley, they identify the role of a kinase called Cla4p in yeast. The kinase is needed to start budding, a form of cell division in which a daughter cell grows out from its parent, before eventually pinching off and going its separate way.

Slot machinery

Other groups, including Peter Schultz and his colleagues at the Scripps Research Institute in La Jolla, California, are using similar chemical strategies to investigate interactions between proteins. The Scripps team has mutated the genes for human growth hormone and its receptor so that the two proteins produced can no longer bind to one another. To do this, they engineered mutations that replaced the amino acids tryptophan and threonine with glycine. These amino acids all have similar backbones, but tryptophan has an additional side-group called an indole. Replacing tryptophan with glycine creates a cavity between hormone and receptor. The researchers reasoned that chemically synthesized indole derivatives would fill this cavity and restore the interaction between hormone and receptor. The strategy worked: Schultz and his colleagues screened a relatively small library of some 200 different compounds and found one small molecule that did the job perfectly⁶. Schultz believes this approach could be used to study the function of a wide range of different protein-protein and protein-nucleic acid interactions.

Meanwhile, Thomas Kodadek of the University of Texas Southwestern Medical



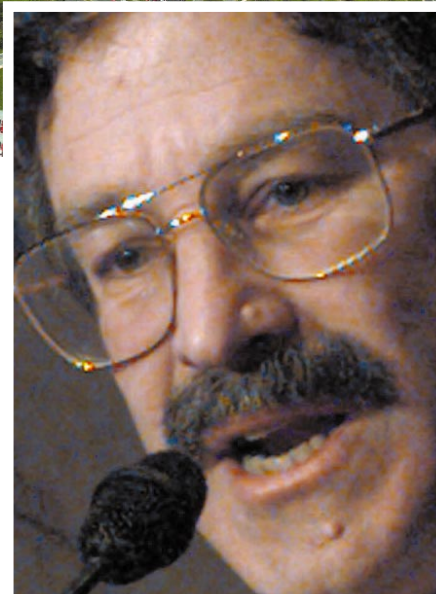
Powerhouse: the National Institutes of Health (above) is backing chemical genetics as a central tool, with Klausner as its champion.

Center at Dallas is looking for small molecules that mimic specific transcription factors. These proteins switch genes on or off by binding to their regulatory sequences. But before transcription factors can do their work, DNA has to be unwound from proteins called histones found inside the cell's chromosomes. So, in addition to transcription factor mimics, Kodadek is looking for small molecules that can trigger this unwinding. The two molecules could then be tethered together. With his colleague Carilee Denison, Kodadek outlined this strategy in a 1998 paper⁷, arguing that it should allow the expression of a "significant fraction" of the genes in the human genome to be controlled artificially. "We are trying to develop a platform technology that solves the problem of how to turn genes on," says Kodadek.

Forward thinking

All of these approaches are similar in concept to established techniques of 'reverse genetics', in which particular proteins are targeted by altering the genes that encode them. But equally valuable are 'forward genetic' approaches, in which researchers mutate the genome at random, screen for mutants with biologically interesting defects, and then work out which gene is involved. This strategy is also being mimicked by chemical genetics. "We're trying to emulate all the principles of genetics using small molecules," says Schreiber.

This effort marks a return to old-fashioned methods of drug discovery. More than



three decades ago, chemists searching for new pharmaceuticals would design synthetic compounds and get biologist colleagues to screen them for useful activity by pumping them into cells and whole organisms. But the development of recombinant DNA technology in the mid-1970s brought in genetically engineered bacteria that could produce pure preparations of proteins identified as potential drug targets. This shifted the initial phases of drug discovery largely to *in vitro* screens, in which drugs companies examined the ability of their libraries of compounds to bind to these proteins.

Today, practitioners of chemical genetics are once again introducing libraries of small molecules into cells. "In a funny way, chemical genetics is a return to the good old days," observes Schreiber. "It's just done with the benefit of modern technology that wasn't available even three years ago." This work could uncover new drug candidates. "There is the potential for an explosion in numbers of those who we define as participating in the

drug discovery process," says NCI director Klausner. But the main focus in many labs is more fundamental biology: the researchers are looking for chemicals with specific effects that might reveal the function of individual genes.

Recent advances in chemical synthesis have facilitated chemical genetics by making it much easier to produce a huge selection of small molecules. Combinatorial chemistry, which builds libraries of compounds by combining smaller molecular building blocks in different configurations, has been used since the early 1990s to generate biologically active molecules⁸. Also, chemists are now developing algorithms that allow them to plan specific synthetic pathways needed to produce a diversity of structurally complex small molecules⁹.

Some labs are also sampling the wares of cash-strapped chemistry institutes from the former Soviet Union, some of which are surviving by selling their huge libraries of compounds to researchers in the West through a company called AsInEx, based in Moscow. "Now it is possible for a biologist to hook up with a chemist's tens or hundreds of thousands of molecules," says Craig Crews, a cell biologist at Yale University in New Haven, Connecticut.

Technologies emerging from genomics are also boosting the field. Robots developed to handle the huge throughput of biological samples needed for genome sequencing can



Schreiber: mimics genetics with small molecules.

be adapted to screen for the effects on cells of thousands of different chemicals. Microarrays, in which libraries of nucleic acids or proteins are immobilized on glass slides, are among the most important tools in functional genomic analysis — and this idea is also being adapted to chemical genetics. Schreiber and his colleagues, for instance, have created microarrays 'printed' with libraries of small molecules¹⁰. Extracts of cells can be added to these chips as a 'pre-screen' in chemical-genetic projects. This allows the researchers to home in on compounds that bind to cellular proteins, reducing the burden of running time-consuming biological assays.

Selection process

Perhaps the most striking demonstration of the power of forward chemical genetics has come from a collaboration between Timothy Mitchison's lab at Harvard Medical School and Schreiber's group at the ICCB. "The whole idea was to find membrane-permeable compounds that affect the proteins required in cell division," explains Thomas Mayer, the postdoc who led the project¹¹. "Then we can use these compounds to the study the function of the proteins."

Mayer and his colleagues began with a library of 16,320 small molecules. If cells are arrested in the early phases of division, they build up high levels of a phosphorylated form of a protein called nucleolin, which can be detected using a light-emitting antibody that binds to it. Using this screen, the Harvard researchers identified 139 small molecules that inhibited cell division. Fifty-three of these targeted a protein called tubulin that forms microtubules — the 'tracks' along which replicated chromosomes move to either end of a cell as it prepares to divide.

Mayer wanted to disable proteins whose role in cell division was unknown, so he

tested the remainder of the small molecules by adding them to mammalian kidney cells and watching the effects on cell division with a microscope after staining them with fluorescently tagged antibodies that bound to both microtubules and chromosomes. One compound, which the researchers dubbed 'monastrol', caused the microtubules to form into a central star-like structure, rather than the usual spindle running from end to end of the cell.

The team went on to discover that monastrol's target is a protein called Eg5. This protein is known to be involved in spindle formation, but researchers do not understand its precise function. With inhibitor and target in hand, Mayer and his colleagues are now adding monastrol to frog cell extracts to dissect the dynamics of spindle formation and the control of cell division¹².

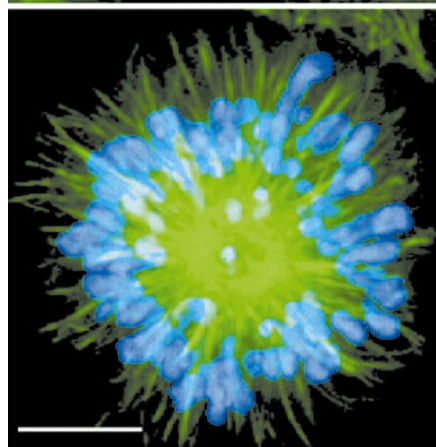
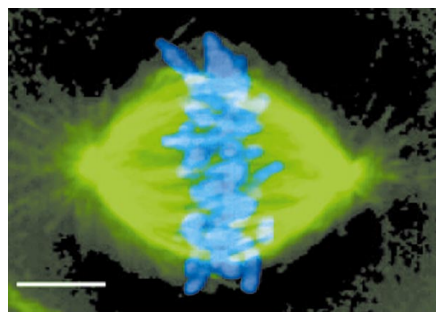
In future, chemical genetics could be applied to developmental studies, for instance in model organisms such as the zebrafish. The advantage of the chemical approach is that it allows genes to be switched off later in an organism's development, says Schreiber. Conventional approaches often induce genetic mutations that are fatal in early development, making it impossible to see any secondary functions a gene might have, which would appear only later on.

If anything is going to hold chemical genetics back, it is most likely to be the cultural difficulty of getting biologists and chemists to collaborate. "It's a tribal clash," says Kodadek. "Some biologists want to hire a bunch of synthetic organic chemists just to make whatever compounds they tell them to." Not surprisingly, the most talented chemists want to work on their own research projects, rather than being at a biologist's beck and call.

Now, as efforts proceed to merge the two disciplines, students are being trained in both fields. "One of the things we are going to observe in the next few years is a broadening of the base of researchers who participate in chemistry," Klausner predicts. If he is correct, these multi-skilled individuals could hold some of the keys to deciphering the human genome, and help to discover tomorrow's life-saving drugs.

Trisha Gura is a freelance writer in Cleveland, Ohio.

1. Bishop, A. C. *et al.* *Nature* **407**, 395–401 (2000).
2. Bishop, A. C. *et al.* *Curr. Biol.* **8**, R257–R266 (1998).
3. Liu, T., Shah, K., Yang, F., Witucki, L. & Shokat, K. M. *Chem. Biol.* **5**, 91–101 (1998).
4. Bishop, A. C. *et al.* *J. Am. Chem. Soc.* **121**, 627–631 (1999).
5. Weiss, E. L., Bishop, A. C., Shokat, K. M. & Drubin, D. G. *Nature Cell Biol.* **2**, 677–685 (2000).
6. Guo, Z., Zhou, D. & Schultz, P. G. *Science* **288**, 2042–2045 (2000).
7. Denison, C. & Kodadek, T. *Chem. Biol.* **5**, R129–R145 (1998).
8. Bunin, B. A. & Ellman, J. A. *J. Am. Chem. Soc.* **114**, 10997–10998 (1992).
9. Schreiber, S. *Science* **287**, 1964–1969 (2000).
10. MacBeath, G., Koehler, A. N. & Schreiber, S. L. *J. Am. Chem. Soc.* **121**, 7967–7968 (1999).
11. Mayer, T. U. *et al.* *Science* **286**, 971–974 (1999).
12. Kapoor, T. M., Mayer, T. U., Coughlin, M. L. & Mitchison, T. J. *J. Cell Biol.* **150**, 975–988 (2000).



Star performer: in normal cell division (top), the chromosomes (blue) move along microtubules (green) in an orderly fashion. But adding monastrol disrupts the system (bottom).

Allowing gene patents could be an expensive mistake for the US

Sir—In the discussion of gene patents¹, it is generally overlooked that these primarily aim at the encoded protein and not the gene. The discoverers try to patent the use of the 'gene product' either as a therapeutic protein, or as an intervention point ('drug target') to treat a disease.

However, although the gene sequence provides necessary information on the protein structure, this information is insufficient, because the mRNA and the original gene product, the protein, are modified in an unpredictable way depending on the temporal and spatial context of their biosynthesis. Who could have predicted the active form of insulin from its gene? Also, the co-receptor of HIV, CCR5, on which a controversial patent has been issued², appears to be a co-receptor for HIV only if it is modified by attachment of sulphate residues³. It is evident that, if a patent covers a 'gene product', it should only give rights on the predicted, unmodified form. Experimental evidence for the existence of the predicted function of the gene product should be provided.

What about patenting gene products as drug targets? The EU patent directive states clearly that the human body and its components in their native environment cannot be covered by a patent. What is a drug target other than a component of the human body in its native environment? This statement can only mean that drug targets are not patentable in Europe, otherwise it would be nothing but a political declamation.

There are other problems with patenting drug targets. Many drugs are still discovered using animal model systems or cellular assays. In the United States, for example, the National Institutes of Health has a number of screening programmes using mice to detect anti-epileptic compounds, and cell cultures to detect cytotoxic compounds, where the molecular target is unknown. If a scientist or a company discovers a drug using such an assay, they would be stupid to try to identify the drug target, because there is some likelihood that the drug target is already covered by a patent. Patenting drug targets is therefore scientifically counter-productive: it enforces a culture of 'not wanting to know'.

Let's assume that drug targets can be patented in the United States but not in Europe. As a consequence drugs could be developed unrestrictedly in Europe but not in the United States. US companies would then be forced to move their drug discovery and development to Europe. Still these drugs could not be used in the United States. As a result patients would have to travel from

the United States to Europe to receive their life-saving drug. What a great perspective for Europe's biotechnology and economy!

Hartmut Michel

Max Planck Institut of Biophysics,
Heinrich-Hoffmann-Straße 7, 60528
Frankfurt am Main, Germany

1. Schiermeier, Q. *Nature* **406**, 111 (2000).

2. Smaglik, P. *Nature* **404**, 322 (2000).

3. Cormier, E. G. *et al. Proc. Natl Acad. Sci. USA* **97**, 5762–5767 (2000).

No room on the carousel for meeting of like minds

Sir—The day has finally arrived when I receive more scientific meeting announcements than junk mail. I read all the programmes with interest, but wonder about the rationale for such a deluge, when information flows freely over the Internet. Notices of meetings and workshops, scientific discussions, the contents of forthcoming journal issues and many other matters can all be found on the Web. So why travel to a meeting when, in most cases, the information is not even new?

As I open the envelopes, my interest fades rapidly when I see the same speakers and topics regularly repeated. I am not criticizing scientists who circle the globe sharing their discoveries and knowledge with the rest of us, but to me they are on a scientific carousel. Each colourful horse is ridden by a speaker with a lecture in his or her right hand. As the merry-go-round rotates, the topic and speaker will cross your path no matter where you are standing.

So what can justify the existence of scientific meetings? They ought to allow us to communicate and share new advances; to discuss new applications for a technique; to hear the latest evidence supporting a discovery; to develop new projects, and so on.

But the whole principle of a scientific meeting — innovation — is ignored in most of these scientific events. The number of meetings inversely correlates with the amount of interesting data to be communicated. On the scientific carousel, a meeting is held for commercial reasons and people attend out of opportunism (local rivalries, for example) or to visit fancy places. In meetings held by scientific societies, however, the spirit of science is maintained, as participants share and discuss advances and challenge each other verbally.

An individual can communicate exciting results on the Web, but the work cannot be generally accepted until it has been challenged and recognized by fellow researchers at meetings. The recognition and respect of one's peers is what constitutes scientific kinship and justifies society meetings. As

for the scientific carousel — like any other fairground attraction, time and market forces will determine its lifespan.

Eleuterio R. Hernandez

Instituto de Bioquímica (CSIC-UCM), Facultad de Farmacia, Universidad Complutense,
28040- Madrid, Spain

Standardized addresses would make web easier

Sir—This may be the age of information, but gaining access to it is often difficult. Finding a site on the World-Wide Web can be like searching for a needle in a haystack.

In India, only half of the 40 laboratories in the Council for Scientific and Industrial Research (CSIR), the country's largest research and development organization, have websites of their own. These 20 institutions follow five different ways of naming their sites. If you don't know the address, you can't reach the site by guessing. None of them end with .gov.in. The Central Mechanical Engineering Research Institute, Durgapur, has opted for www.cmeri.com. Some end with .org — which gives the impression that the pages are hosted by a service provider located in the United States, where web addresses do not end in a domain indicative of the country. Most web-based search engines are unable to help, since they only reach some 16% of the web (see *Nature* **400**, 107–109; 1999).

Adopting uniformity in naming domains could circumvent this problem. Many institutions around the world use their name or acronym as the lowest domain and move up hierarchically. For example, all the institutions in the CSIR could use .csir.dsir.gov.in, since CSIR is part of the Ministry of Science and Technology's Department of Scientific and Industrial Research. The Centre for Cellular and Molecular Biology could be at www.ccmb.csir.dsir.gov.in. Educational institutions should follow the same convention. By understanding the organizational hierarchy of any institution one should be able to guess the web address.

The next improvements could be to the content, organization and facilities provided by the web pages. In the scientific world, concerted efforts are under way to realise the full potential of electronic communication, with e-print archives, online journals and the development of E-Biomed. Web pages for government-funded research and development laboratories need to offer more than just basic facts about the organization. They should show depth of content and should include details of their scientific activities, search facilities and routine updates.

Sasidharan Rajkumar

Nagarjuna Group, Nagarjuna Hills, Hyderabad 500 082, India

The Durban Declaration is not accepted by all

Sir — In response to recent action by President Thabo Mbeki of South Africa and in advance of the International Conference on HIV/AIDS held in Durban on 9–14 July, the Durban Declaration¹ was prepared by a committee representing a consensus of “181 scientists and front line physicians”. Before publication in *Nature*, it was circulated: “To get as many names of scientists and doctors to sign on. Names of signatories will appear on the *Nature* website. If you would like to sign on, we would be delighted. Send me an e-mail confirming this. To economize space on the website, we have to name people in a single line. Many of you will say that HIV/AIDS is not your area. However, over the years you have heard enough of the arguments to understand the association. Furthermore, many of you know well infectious diseases and understand Koch’s postulates. If you have colleagues in the laboratory or in the clinic who you feel would like to sign, please ask them. The more the better. However, please note that in order to be authoritative we feel it necessary to restrict the list to those with major university qualifications.” This is an extract from the circular distributed on behalf of the organizing committee which included Luc Montagnier, Catherine Wilfert, David Baltimore, Sir Aaron Klug (as President of the UK Royal Society), and many other well-known names and organizations from developing countries as well as from the West.

Briefly, the authors of the declaration state that AIDS/HIV is spreading as a pandemic now affecting 34 million people, of whom 24 million are in sub-Saharan Africa. They say the disease began there as a viral infection of chimpanzees and monkeys conveyed somehow to humans, and is now spreading worldwide by heterosexual and mother-to-infant transmission. The authors consider that their evidence supporting this hypothesis is “clear-cut, exhaustive and unambiguous”; that most people with these infections will develop AIDS within 5–10 years unless treated; and that “there is no end in sight” until research based on their hypothesis leads to a vaccine to supplement safe sex, health education and other, simpler approaches to avoidance and prevention.

With no end in sight after 17 or more years of intensive research, priorities and incentives, one might think that this consensus would be open to alternative approaches, but the authors of the declaration are emphatic that this is not needed because the evidence that HIV is the cause of AIDS has met or exceeded the “highest standards of science”. By implication, any other evidence is therefore a deception,

even less likely to lead to a successful vaccine, curative drug or hypothesis.

Our objection to the Durban Declaration is factual and verifiable from data published in the early 1980s (refs 2–4). We believe that World Health Organization (WHO) figures produced since then⁵ can be interpreted to say that AIDS first appeared and spread, not in Africa but in US urban clusters of mainly white, affluent, promiscuous homosexual men and drug addicts, and then spread, on a lesser scale, in Europe and Australasia but hardly at all in Asia. Disastrous epidemics due to heterosexual transmission of HIV were confidently predicted in general populations of developed countries⁶ but they never happened. AIDS has diminished in incidence and severity though it is continuing in female partners of bisexual men and some other communities engaging in or subjected to behaviours which carry high risks of infections, various assaults and misuse of drugs.

In sub-Saharan Africa, AIDS was reported later^{7,8}, with an alarming frequency in mothers and infants not seen in the United States or Europe. Sentinel surveillance by the WHO shows correlation between this frequency and the seroprevalence of HIV, but there are unmeasured overlaps with other major diseases and deprivations which, together with anomalies in classification, distribution, transmission and country-specific pathogenesis, and especially cross-reactions in serological tests^{6–9}, raise questions about the accuracy of diagnosis and approaches to control.

In the absence of satisfactory, or of any, answers from the consensus to his specific questions on this matter, President Mbeki invited us to join other experts with differing viewpoints in a panel to explore the way forward to control AIDS in Africa. Unlike the signatories to the Durban Declaration, we claim no exhaustive and unambiguous unanimity. There are differences between ourselves and with other panellists, and we are happy to acknowledge possible convergence with certain priorities favoured by the declaration’s authors. But we reject as outrageous their attempt to outlaw open discussion of alternative viewpoints, because this reveals an intolerance which has no place in any branch of science. Our viewpoints could also explain the failure to prevent the spread of AIDS in high-risk populations in the West, amounting, in the United States now, to almost 700,000 registrations — an unbeaten score in the global tally of this disease.

Gordon T. Stewart, MD

3 Lexden Terrace, Tenby, Pembrokeshire SA70 7BJ, UK (Emeritus Professor of Public Health, University of Glasgow)

Other signatories to this letter; full addresses available from G.T.S.

Sam Mhlongo, MB, BS Professor of Medicine, MEDUNSA, Johannesburg, South Africa

Etienne de Harven, MD Emeritus Professor of Pathology, University of Toronto, Canada

Christian Fiala, MD Obstetrician, Vienna, Austria

Claus Kohnlein, MD Physician, Städtisches Krankenhaus, Kiel, Germany

Andrew Herxheimer, MD Pharmacologist, London, UK

Peter Duesberg, PhD Professor of Molecular Biology, University of California at Berkeley, USA

David Rasnick, PhD Research Fellow, Dept of Molecular Biology, do

Roberto Giraldo, MD Physician, New York City

Manu Kothari, MD Pathologist, Seth GS Medical College, Bombay, India

Harvey Bialy, PhD Research Scholar, National University, Mexico City, Mexico

Charles Gesheker Professor of African Studies, California State University, Chico, California

1. Durban Declaration, *Nature* **406**, 15–16 (2000).

2. *Morbidity Mortality Weekly Reports* **30**, 250 (US CDC, Atlanta, 1981).

3. *Morbidity Mortality Weekly Reports: Update on Acquired Immune Deficiency Syndrome (AIDS)*, USA **31**, 507–514 (1981).

4. Gottlieb, M. S. et al. *N. Eng. Med. J.* **305**, 1425–31 (1982).

5. *Weekly Epidemiological Records* (WHO, Geneva, 1981–2000).

6. Cox, D., Anderson, R. M., Hillier, H. C. (eds) *Phil. Trans. R. Soc.* **325**, 37–187 (1989).

7. *International Classification of Diseases*, 10th revision (WHO, Geneva, 1992).

8. Root-Bernstein, R. *Rethinking AIDS* (MacMillan, New York, 1993).

9. Kashala, O., et al. *J. Inf. Dis.* **109**, 296–304 (1994).

Sky was not the limit for music-loving Herschel

Sir — In a review of Ludmilla Jordanova’s *Defining Features: Scientific and Medical Portraits 1660–2000*, Lisa Jardine describes a portrait of the astronomer William Herschel that includes the night sky¹. Commenting on portraits, she writes: “As in Herschel’s case, what distinguishes the scientist from other celebrities ... is that, where the politician might hold a pen and the musician his instrument, the scientist’s portrait is likely to allude to the scientific breakthrough itself”.

From this comment, one might not guess that, before he became a famous astronomer, Herschel was a professional musician of some distinction. He was active as an oboist, a violinist, an organist, an orchestra director and — most interestingly — a composer². Herschel wrote symphonies, concertos, chamber music and vocal works^{2,3}.

In June 1792, Joseph Haydn visited Herschel’s observatory near Slough⁴. By that time, Herschel was a full-time astronomer, and it appears that Haydn’s visit was motivated primarily by an interest in Herschel’s astronomical activities.

Martin F. Heyworth

VA Medical Center and University of Pennsylvania, University & Woodland Avenues, Philadelphia, Pennsylvania 19104, USA

1. Jardine, L. *Nature* **405**, 397–398 (2000).

2. Cudworth, C. & Jeans, S. “Herschel, Sir William” in *The New Grove Dictionary of Music and Musicians* (ed Sadie, S.) vol. 8, 520–522 (Macmillan, London, 1980).

3. *The Oboe Concertos of Sir William Herschel* (ed Jerome, W. D.): *Memoirs of the American Philosophical Society*, Vol. 225 (American Philosophical Society, Philadelphia, 1998).

4. Landon, H. C. R. *Haydn in England 1791–1795 (Haydn: Chronicle and Works, Vol. 3)* 176–177 (Thames and Hudson, London, 1976).

Good and bad science in US schools

One-third of US states have unsatisfactory standards for teaching evolution.

Lawrence S. Lerner

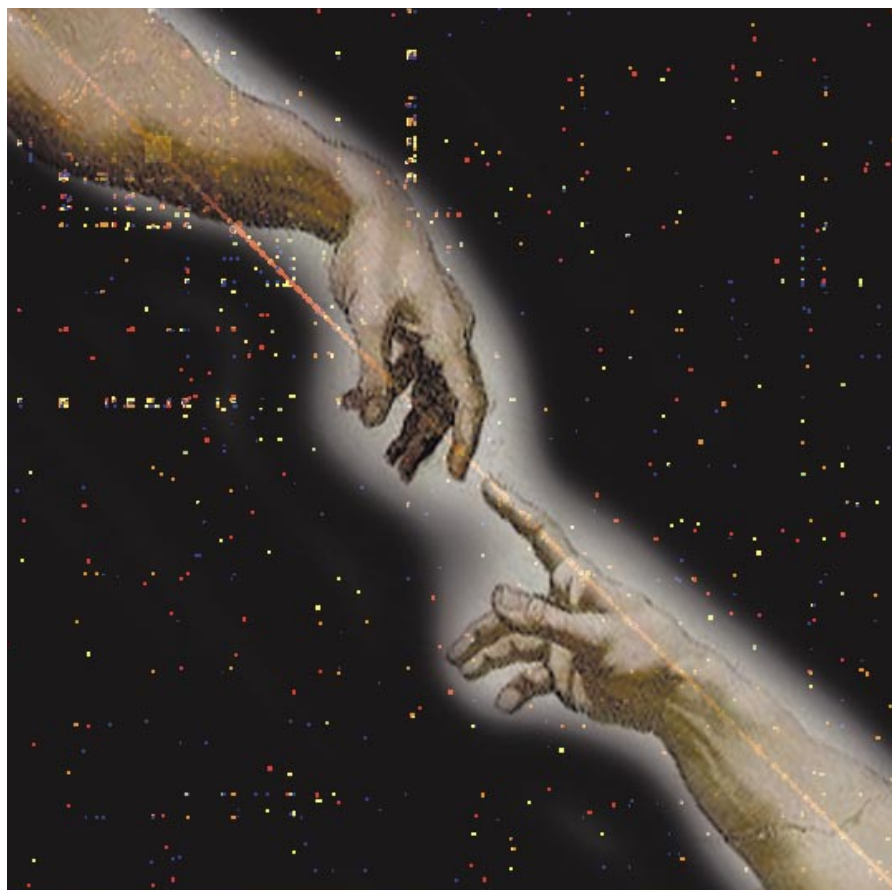
Almost all science is the study of the evolution of systems in time. Biology is no exception; its central organizing principle is the evolution of living things, just as geology centres on the evolution of the Earth and astronomy on the evolution of the Universe.

That evolution is the central organizing principle of all the historical sciences is not a controversial issue among scientists, or among most educated people. Consequently, science teaching worldwide treats evolution as routine. The United States is the exception. In much of the country, teaching evolution as scientists see it — particularly biological evolution — to K–12 (age 5–18) students still evokes bitter controversy, 75 years after a landmark court trial first brought the matter to broad public attention. A variety of organizations and people object to the teaching of the facts and theory of evolution in public schools at the primary and secondary level.

This controversy is not really about science but about religion and politics. Objectors assert that evolution has not taken place, that scientists' picture of the Universe is misguided, that it is 'only fair' to present creationist views to students in tandem with evolution, even that teaching evolution will lead children into immoral lives. The first two assertions are part of 'creation science', a pseudoscientific rival to evolution that US courts have repeatedly ruled to be thinly veiled religion.

This essentially non-scientific controversy is reflected in the primary–secondary science standards of many US states. Objections to the teaching of evolution are expressed in two main ways. First, the fundamental concepts and facts of evolution are covered to some extent, but the word 'evolution' is replaced, in the context of biology, by incorrect and misleading euphemisms such as 'change over time'. Second, the subject is avoided altogether or barely mentioned, reducing the sciences, especially the biological sciences, to disjointed lists of facts.

I have assigned each of the states a grade for its treatment of evolution* (see Table 1). On balance, the news is good. Thirty-one states (almost two-thirds) do at least a satisfactory job of dealing with the central organizing principle of the historical sciences at



The way in which biology is taught varies more in quality and quantity than for any other scientific subject.

the level of their statewide academic standards. Ten states do a very good to excellent job (grade A) of presenting evolution and 21 do a good or satisfactory job.

The bad news is that more than a third of all the states are less than satisfactory, and thus seriously damage or even erase the possibility of teaching science to young people as more than a confusing collection of facts. Six states rate an unsatisfactory D and 13 an F or worse, their standards being essentially useless for purposes of teaching evolution.

Why are standards important?

Statewide academic standards are meant to be the foundation for a host of curricular

activities that affect what goes on in a state's classrooms: as the basis for teachers' lesson plans, statewide examinations and textbooks; as a method for parents to assess their children's progress; and as a basis for uniformity and ease of transition for students moving to new schools, for example.

I appraised state science standards with respect to their overall quality in *State Science Standards: An Appraisal of Science Standards in 36 States* (March 1998) and *The State of State Standards 2000* (January 2000), both published by the Thomas B. Fordham Foundation (available at <http://www.edexcellence.net/fordham/foreports.html>). I used 25 criteria in five categories to evaluate quality, each standards document being considered for its purpose, expectations and audience;

Table 1 **Distribution of grades**

Grade	Number of states	Treatment of evolution
A	10	Excellent or very good
B	14	Good
C	7	Satisfactory
D	6	Unsatisfactory
F	12	Useless or absent
F–	1	Disgraceful

* This Commentary is a version of a longer report to be published on 26 September which contains full details of the analysis and criteria for the rankings of states described here. The full report can be obtained from the Thomas B. Fordham Foundation in Washington DC at 202-223-5452 or from www.edexcellence.net. It will also be presented at a public AAAS symposium, 'The Teaching of Evolution in US Schools: Where Politics, Religion, and Science Converge', on 26 September.

organization; coverage and content; quality; and negative elements detracting from the standards. Evolution was but one of many topics considered in these appraisals; here I provide a detailed assessment of evolution teaching in US schools using a separate set of criteria which are explained briefly below and in greater detail in the full report.

The peculiar place of evolution

What do we mean by evolution, and what is its place in the sciences, since almost all science is the study of the evolution of one system or another? So evolution is an indispensable concept across all the sciences. But in the United States (and almost nowhere else), the public attention attracted by biological evolution is different in kind and intensity from that attracted by evolution in other scientific fields. As a result, the way in which biology is taught varies considerably more in quality and quantity than for any other scientific subject, with a significant spillover into geology, the evolution of the Solar System and cosmology. The physical sciences are affected as well, but more indirectly.

The following questions seem to me central. What constitutes good academic standards as they concern evolution, biological and otherwise? What kinds of public pressures oppose good standards in this area? What are the bases for the opposition? How have the states reacted to these pressures? What is the effect on the quality of science standards overall — and thus on the quality of science education in the United States?

Any decent education in science requires that the student comes to understand the central role of theory in scientific methodology. This understanding does not emerge full-blown — young people grow into the ability to understand the abstractions essential to the methodology (see box, below). Lacking this insight, the student inevitably comes to see the sciences as a stultifying heap of disconnected facts, some of them counter-intuitive and all hard to sort out. This luckless student soon learns how to commit the required facts to short-term memory, squeak past the next test, and then forget what he or she has so painfully memorized. The present state of scientific literacy among US adults

bears dour witness to the ubiquity of this kind of learning experience.

Enormous amounts of data of extraordinarily diverse kinds have been interrelated and made understandable on the basis of the theory of biological evolution. Today, that theory informs, and is vindicated by, a larger and more varied body of evidence than that associated with any of the other major branches of science. Moreover, biological evolution is seamlessly joined with geological evolution, and is completely consistent with the principles of physics and chemistry.

Extrascientific issues

Anti-evolution views span a wide spectrum. Moreover, they have evolved over time, responding to judicial and social pressures, competing to fill 'ecological' niches, and scouting for new ones to occupy. The literature devoted to anti-evolution views and their refutation is diverse and vast.

In contrast to the situation in K–12 schools, no controversy exists at the university level, where curricula are more or less fully under the faculty's control. As science faculty members are experts in their fields, they share a consensus as to basics. In contrast, K–12 instruction is subject to considerable intervention from people such as school-board members and legislators with no expertise in, and often little or no knowledge of, the fields whose curricula they govern, and hence who are vulnerable to irrelevant lobbying and pressures.

Anti-evolutionists tend to object to three main premises. First, to achieve the diversity of life observed today, the evolutionary process has required several billion years. Second, all living things, including humans, are descended from common ancestors. And third, the evolutionary process is a natural one susceptible to scientific investigation and so by definition cannot include supernatural intervention as a necessary component.

The first of these premises conflicts with a particular interpretation of the start of the Book of Genesis and of Old Testament genealogy, according to which the Universe is a few thousand years old rather than tens of billions of years as the scientific evidence implies. This particular interpretation of the Bible, generally called young-Earth creationism, is held mainly by a subset of evangelical Protestants and some ultra-orthodox Jews and Muslims. Young-Earth creationists fear that the alternative interpretations of Genesis supported by most Christians and Jews undermine the entire authority of the Bible. Other religious groups object for the opposite reason. In particular, members of the African-American sect the Nation of Islam hold that the Universe is trillions of years old, whereas adherents of some Native American religions hold that their ancestors have been located in their traditional tribal areas for ever (that is, for an infinite time).

Essentials of good standards

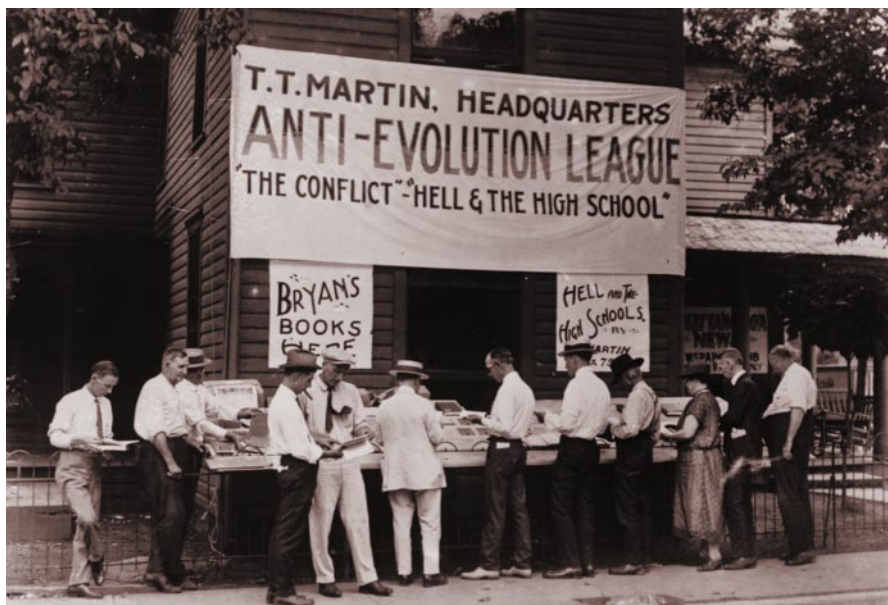
What is needed to provide the student with a good understanding of the life sciences and the unifying role of evolution? At the primary grade levels, standards should focus on those basic facts and ideas of evolution that can later be incorporated into broader world views. At the K–3 level (age 5–8), students should be expected to understand that all living things reproduce; that offspring are similar to but not exactly like their parents; that offspring have to grow up (or change) before reproducing themselves; that there is a fit between individuals, or species, and their environment (for example, terrestrial, aquatic, aerial); and that the Earth is more than four billion years old, allowing much time for biological as well as geological evolution.

At higher grade levels (ages 9–12) these ideas can be developed into an understanding of the nature of competition for survival between and within species; the consequence that not all offspring live long enough to reproduce; the limitation imposed on the number of offspring that survive by such environmental factors as availability of food and water, predators and temperature; the variability among individuals that leads to differential survival in a particular environment; the specialization of species to fit ecological niches and the impact of environmental change on the tenability of those niches; the underlying role of genetic variation that results from sexual reproduction and random mutation; and the non-random way that natural selection operates on the existing population despite the many random factors that determine the survival of any individual.

At the middle- and high-school levels (age 11–18) these ideas can be unified, and such concepts as genetic drift, sexual selection and other significant mechanisms can be introduced. Coevolution and the complex interactions of ecosystems are important applications of the basic concepts. The magnitude of the geological/evolutionary timescale is so different from the timescales of everyday life that it is difficult to grasp, and must be introduced with care. The fact that the same general timescale underlies both geological and biological evolution is an important link between the two sciences.

In parallel to these macroscopic concepts, the underlying microscopic mechanisms must be introduced at suitable grade levels. These include the relation of genotype to phenotype, DNA as an information carrier, the expression of DNA in protein synthesis and the implications thereof at the various levels of organization from organelles through cells, tissues, organs and individual organisms, to populations.

It is also important to teach that biological evolution does not take place in a vacuum. The biota and the non-living parts of the Earth coexist with and influence each other. Therefore, the facts and, subsequently, the theoretical structure of geological evolution must be introduced in parallel with biological evolution. Similarly, the Earth is part of the Solar System and the Solar System part of a hierarchy of still larger structures, up to the Universe as a whole. The student should be empowered to view the history of the Universe, from the general cosmological picture down to the smaller scales characterizing the Earth and its smaller elements, in a unified perspective.



Books for burning? Dayton in 1925, when John Scopes was on trial for teaching the theory of evolution.

The second premise is objected to by young-Earthers and others who hold that humankind has a special, divinely ordained place in the Universe and is the central concern of the Divinity. According to this belief, God could not have lumped humans with mere animals, let alone other living things. Hence, teaching the evolutionary relationship of humans to other animals inevitably undermines any possible moral or ethical teaching. If it is taught that humans are 'only animals' they will 'act like animals'. This, it is alleged, will lead to atheism, communism, socialism, Nazism, inflation, homosexuality, women's liberation, sex education, teenage sex, abortion, pornography, family breakdown, school shootings, crime, alcoholism and drug addiction — to name but a few. The same believers also often suggest that certain political and religious positions presuppose adherence to creationism; that is, a person cannot truly be a religious or political conservative without also being a creationist.

The third premise, although shared by the groups discussed above, is the special province of a class of anti-evolutionists called intelligent-design or irreducible-complexity advocates, who have revived the position of John Ray in the seventeenth century and of William Paley in the early 1800s. Intelligent-design advocates believe that adherence to an evolutionary view of the biosphere is conducive to atheism, or even that only atheism is consistent with an evolutionary view of the Universe, to which they assign the name 'naturalism', a term they construe as pejorative. They say that living beings are too complicated to have evolved, and that their creation by an intelligent (read divine) designer is just the entrée into the natural world that God requires for general belief.

Evolution has had its equally vehement opponents on the political left. The classical

marxist view of 'socialist man' was official doctrine in Stalin's Soviet Union. Holding that the evils of society and human immorality stem exclusively from socioeconomic injustices, and that humans will become entirely virtuous in the Marxist utopia, the Stalinist view required rejection of any implications that human behaviour might have biological roots. Utopian socialists, although their approach was more benign, held similar views on the perfectability of human nature under the proper socioeconomic conditions. Although Stalinism is extinct as a political power, there still exists in the United States a small but sometimes vocal left-wing intellectual opposition to evolution.

Why anti-evolutionism persists

Not surprisingly, the states wrestling with the teaching of evolution are largely those that have substantial populations of Protestant evangelicals. Most evangelicals are probably not anti-evolutionists, but those who are comprise the largest, most significant bloc opposing the teaching of evolution in public schools. The success of creationist efforts at the state level is probably due to a sympathetic chord struck by creationist activists in a wider public who do not share, or lack strong interest in, creationist ideologies.

A significant number of Americans who are not scientists feel uncomfortable about evolution in general and biological evolution in particular. Polls suggest that about a quarter of Americans adhere to the literal biblical account, perhaps a third are strongly convinced of the validity of the scientific account, and the rest are ambivalent.

A recent poll surveyed US attitudes towards the teaching of evolution and creationism in the public schools. It was a representative, national survey of 1,500 Americans conducted by telephone by DYG in Novem-

ber 1999. Two-thirds of the respondents believed that only evolution should be taught as science. Of these, 20% held that only evolution should be taught at all; 17% that evolution should be taught in science classes while religious creation should be taught in other classes; and 29% that evolution should be taught in science classes as a scientific theory with creationism mentioned as a belief. The 'equal time' view, that evolution and creationism should be taught together as competing views in science classes, was held by 13% of all the respondents, with 4% more feeling that both should be taught but uncertain as to the details. Finally, 16% believed that only creationism should be taught. Only 1% belonged to the 'don't know' category. Interestingly, given the wide range of evolution teaching among states, these percentages varied less than 10% from region to region.

Evolution of anti-evolution pressures

Holding that biological evolution is in conflict with religion and conducive to immorality, anti-evolutionists have, unsurprisingly, tried to interdict its teaching in the classroom. The only anti-evolutionists to have had significant influence on K–12 public education so far have been the young-Earth creationists. A long series of court decisions, including two by the US Supreme Court, have held that the underlying premises of creationism, however they are presented, are religious rather than scientific and so have no place in the public-school science classroom.

Largely in response to these decisions, creationism has itself evolved. When *Epperson v. Arkansas* struck down laws forbidding the teaching of evolution in 1968, creationists responded with two curiously conflicting claims: first that creationism and evolution are both based on articles of faith and are therefore both religion, not science; and second that creationism is a scientific approach to the study of the living world just as is evolution.

The second of these claims became the basis for model legislation introduced in many state legislatures and passed in at least two. This 'balanced-treatment' legislation required that whenever 'evolution science' was taught, 'creation science' had to be taught as an equally valid, competing explanation of the living world. It was further argued that 'scientific creationism' was independent of but compatible with 'biblical creationism'. A series of court decisions, culminating in a 1987 Supreme Court ruling, declared that this argument, too, was an attempt to veil religion as science. In particular, creation implies a creator, which is a religious concept. For this and related reasons, the courts held that 'creation science' was not science and had no place in the science classroom, although creation stories could certainly take their place in other studies.

The failure of 'creation science' to gain

acceptance by the US Supreme Court led to the re-emergence of the 'intelligent-design' creationists. These advocates do not necessarily object to an old Universe, or even to a certain amount of evolution. But they do insist on supernatural intervention as necessary to any explanation of the living world. Although they try to minimize their differences with the competing classes of creationists for political reasons, their approach is vehemently opposed by the young-Earthers.

Response to creationist pressures

States that respond to these various anti-evolution pressures do so in different ways and degrees, usually in one or more of the following forms:

- Standards include many of the central principles of evolution — usually briefly — but the word evolution is avoided. Inaccurate and misleading euphemisms such as 'change over time' are used instead of the 'E-word'.
- Biological evolution is ignored. Geological evolution, the history of the Solar System and cosmology may be treated, often even using the word evolution. Fossils are sometimes mentioned, but only in the context of geology, not biology.
- Evolution of plants and animals is treated to some degree but human evolution is ignored.
- All scientific discussions that imply an old Earth or Universe are deleted.
- Creationist jargon is used.
- All textbooks must carry a disclaimer that calls evolution controversial or labels it a theory, not a fact, misusing these terms in their everyday rather than their scientific senses.
- Some or all of the historical sciences are treated lightly but no attempt is made to elucidate the connections among them.

The grades I give here (see Table 1) for the treatment of evolution in state science standards reflect the extent to which states have resorted to these anti-evolution tactics. I give states lower marks if, for example, they avoid mention of the word evolution, they ignore human or biological evolution, or they use creationist jargon.

Evaluation of state standards

Forty-nine of the 50 states and the District of Columbia have published science standards. As a matter of policy, Iowa does not write statewide academic standards in any subject. The treatments of evolution in the 50 science standards fall into six categories, which I denote A, B, C, D, F (and F-) to correspond with traditional letter grades. Of the 31 states that have satisfactory-to-excellent treatments of biological evolution, only nine treat human evolution explicitly and another 11 by implication; the rest do not cover human evolution at all.

Of the 19 states in the less-than-satisfactory ranks, 10 cripple their treatment of evolution through seditious avoidance of the 'E-word', one state uses the word only once,

Anti-evolution views have evolved over time, responding to judicial and social pressures.

and one state hides it. Of these 12 states, eight attempt to teach some evolution, but do a poor-to-awful job of dealing with the subject. The seven remaining states that receive less than satisfactory grades do mention the word evolution but do not do a good job of covering the concept.

Of the 19 states receiving unsatisfactory grades, three ignore the topic of biological evolution altogether, and one not only shuns biological evolution, it also deletes all references, direct or indirect, to the age of the Earth or the Universe, even including radioactive decay (see *Nature* 406, 552; 2000).

The US public education system is based on the concept of local control with limited oversight from the state. Today, in most of the country, there is a strong public consensus for quality control at the state level, and all states have a variety of mechanisms for accomplishing this, such as curriculum standards, statewide examinations, teacher certification, class-size rules and compulsory education laws. The fact that state subventions are important in the financing of school districts is, of course, a significant factor in underpinning a modicum of control. So one can characterize US public-school governance as a modified diffuse democracy.

Democratic decisions are normally made by majority rule. We believe, of course, that a well-educated electorate is an essential basis for workable democracy; this has always been a cardinal argument for public education. But democracy cannot permeate all aspects of every social institution. It is certainly not consistent with the education process itself because the teacher directly supervises the progress of the students, using his or her superior knowledge and adhering to standards imposed from levels above the classroom.

In a democratic society, citizens who do not like the existing state of affairs can change it. But nature is not so flexible. We may find Newton's second law of motion contrary to common sense because it links acceleration, rather than velocity, to applied force. But we cannot change this fact; what we can do is learn how to manipulate it. We may believe it an affront to human dignity that the Earth is not at the centre of the Universe but we cannot move it there.

Science is also undemocratic in the social sense that those who do not have the scientist's special knowledge, skills and experience

cannot have equal voice in achieving a scientific consensus concerning a class of phenomena. The public school has no authority to impose opinions on its students, but it has the duty to explain to them the consensus of scientists on any particular issue, and the methodology by which scientists proceed to discover new knowledge and merge it into that consensus.

Biological evolution is one of the most important of many broad issues on which almost all working scientists agree. There may be a few people with scientific credentials who disagree, but they do not contribute to the progress that is the hallmark of science. Analogously, there are a few scientists who do not believe that HIV is the cause of AIDS, but they have contributed nothing to the development of the antiviral drugs that have so greatly improved the prognosis for patients over the past decade or so. It is not simply that these dissenters are wrong, because wrong answers can sometimes stimulate controversy that helps lead to correct answers. Rather, as the physicist Wolfgang Pauli liked to say, they are "not even wrong". That is, their arguments are useless and even detrimental to the pursuit of further knowledge.

This being the case, the publication and maintenance of scientifically accurate academic standards is a vital quality-control function of the states. Given the far-reaching ramifications of evolution in the life sciences as well as in other sciences, a complete and proper exposition of evolution is an essential constituent of state science standards. Short-changing, distorting or omitting evolution harms the teaching of the sciences, and makes it difficult for the student to come to a clear understanding of how science works.

Given this state of affairs, a school district or a state cannot argue that it is a simple matter of democracy to advocate a scientifically unacceptable opinion because a majority or vocal minority of citizens holds that opinion. One can understand the desire of parents to raise their children to think as they do. But if the parents have a poor understanding of the content and methods of science, it is well if they hope and expect that their children will understand science better than they do. In doing so, parents will provide the means to expose their children to expertise beyond their own. Indeed, that is why most parents want to send their children to school.

About two-thirds of US states today have science standards that are consistent with this educational philosophy. We owe it to the students in the other third to raise the standards to the same level in these states. Poor education, wherever it may be and in whatever discipline, affects us all. ■

Lawrence S. Lerner is emeritus professor of physics and astronomy at California State University, Long Beach, Long Beach, California 90840, USA.

e-mail: lslerner@csulb.edu

Is your journal really necessary?

Science may best prosper if print journals are replaced by online communities.

Declan Butler

"What I like about *Nature* is that it is the sort of journal you read in the toilet." While there is no doubt a quip to be made about other potential uses of *Nature's* pages in such a setting, this comment, by a prominent French geneticist, is flattering. It testifies to reader appreciation of the costly professional input that goes into producing readable journals of high scientific and editorial quality.

Many of the challenges faced by the staff of *Nature* are shared by the editors of the dozen or so fledgling journals reviewed in the following pages. The expressed motivations for creating these journals encapsulate many core functions of journals: to attract top papers, to promote rigorous peer review, to collate material within a discipline — in the words of *Traffic: The International Journal of Intracellular Transport* (see page 299), to create "a central journal to gather together publications that are of most interest to those working on cellular trafficking" — or work scattered across many disciplines, as in the case of *Interfaces and Free Boundaries: Modeling, Analysis and Computation* (see page 297).

Another function, well exemplified in three new gene-therapy journals (see page 292), is to review, to regularly take stock of published work, and trends. All these core functions of a journal share a commitment to impose intellectual rigour and high editorial standards on an exponentially increasing body of knowledge, so as to make that information more accessible and placed in a wider context.

Faced with such needs, the response has usually been to create printed journals to make the new data more manageable. But are there now better ways of meeting the same needs? The answer is an unequivocal 'Yes'. Admittedly, print has a look and feel missing — for the moment — in the electronic world. But the plethora of print journals is doomed to extinction; it makes no economic sense and is increasingly a hindrance to science itself.

Not all print journals will disappear. Journals whose content can command a large readership will continue to exist, and indeed flourish, in print, as their economics are akin to those of the magazine market. But the bulk of journals are consulted no more than 50 times a year in a typical library, and only 15% more than 250 times. Subscribing just to the handful of journals reviewed here will set you back several thousand dollars. The costs of print are difficult to justify for most journals (see *Nature* 397, 195–200; 1999). In a free

market, high-cost/low-circulation journals would be forced to go electronic, or disappear.

The current proliferation of low-circulation journals means that even the richest libraries cannot subscribe to them all. The toll-gates of high subscriptions of thousands of print journals are thus a fundamental and unnecessary barrier to knowledge. But this so-called "serials crisis", although immediately acute, is not the major reason most journals are destined to exist only electronically.

What will drive change is Internet functionality. I do not refer to small improvements, such as better and faster searching, but to an imminent paradigm shift in scientific publishing and data handling. The core functions of a journal are the intelligent commissioning, editing and grouping of material to meet the needs of communities.

But in the future an electronic paper's content will be linked and matched automatically to related material across the entire literature, using increasingly sophisticated algorithms, rejuvenating serendipity and interdisciplinarity. The web is ideal for aiding the core journal function of regrouping work scattered across many disciplines.

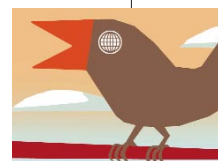
But in this inevitable electronic world, the invisible hand of professional publishers and editors will be as important as it is now, and there will be costs attached. Indeed, as the flood of information grows, more and not less human editorial skill will be needed to make sense of it. At the same time, market demand for access across the entire literature will drive imaginative deals between publishers and libraries to make such access affordable. Transferring these costs from libraries to the users themselves, for example on the basis of metered use, would also allow market forces to turn the distorted scholarly publishing market into a competitive one.

The possibilities of sophisticated matching of personalized editorial selections across large swathes of the literature, and the need to lower barriers to access, should in themselves be sufficient to convince scientists tempted to create low-circulation print journals to consider web-only options. Arguments that electronic-only will hinder access of developing countries to science is nonsense. The reality is

that a library in Kinshasa would be lucky if it could afford to subscribe to a handful of print journals; the web promises developing countries access to scientific information they could previously only have dreamed of.

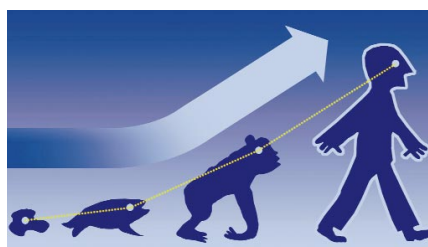
But the essential function of a journal is to serve a particular community. The next web revolution will be a plethora of next-generation communities linking papers, people and data. So next time you think about launching a print journal, unless you have sufficient readership to survive in a free competitive market, do your colleagues and science a favour by considering instead what your community needs, and launch the answer online. I predict that this change will occur in under five years; if I am wrong, I will eat my journal.

Declan Butler is European correspondent of *Nature*.



New journals index

- p292 ● Molecular Therapy
- The Journal of Gene Medicine
- Genetics in Medicine
- Genesis: The Journal of Genetics and Development
- p293 ● The International Journal of Neuropsychopharmacology
- Pituitary
- p294 ● Animal Conservation
- Journal of Environmental Monitoring
- Journal of Environmental Policy & Planning
- p295 ● Wind Energy
- Green Chemistry
- p296 ● Foundations of Chemistry: Philosophical, Historical, Educational and Interdisciplinary Studies of Chemistry
- Crystal Engineering
- p297 ● Interfaces and Free Boundaries: Modeling, Analysis and Computation
- The European Physical Journal E – Soft Matter
- Nonlinear Phenomena in Complex Systems
- p298 ● Journal of Nanoparticle Research
- International Journal of Phytoremediation
- p299 ● Traffic: The International Journal of Intracellular Transport
- Evolution & Development
- p300 ● Drug Resistance Updates



Healing in the genes

Molecular Therapy (American Society of Gene Therapy)

editor-in-chief Inder E. Verma
Academic. 12/yr. \$245 (US, Canada), \$305 (elsewhere), \$150 (ASGT member), \$50 (ASGT associate member)

The Journal of Gene Medicine (European Society of Gene Therapy)

editors Olivier Danos, Kay Davies, Pierre Lehn & Richard Mulligan
Wiley. 6/yr. \$380 (institutional), \$265 (individual), \$165 (ESGT member)

Genetics in Medicine

editor-in-chief Richard A. King
Lippincott Williams & Wilkins. 6/yr. Institutional: \$439 (US), \$474 (elsewhere); individual: \$219 (US), \$254 (elsewhere); in-training: \$89 (US), \$124 (elsewhere)

Karin Sitte and Robert Williamson

Gene therapy became a buzzword of genomics in the 1990s, and is still a hot and sometimes controversial topic in medicine and the popular press. In spite of few clinical successes to date, and some widely reported dramatic failures, gene therapy seems to be as topical as ever in the new millennium. The emergence of two new journals in gene therapy, *The Journal of Gene Medicine* and *Molecular Therapy*, demonstrate that although the two established journals in the field, *Human Gene Therapy* and *Gene Therapy*, attract specialist manuscripts of high quality, there is room for competition in the field.

The new journals have been sponsored by gene-therapy societies (respectively in Europe and the United States) to communicate information faster and in a way claimed to be lacking in the existing journals. These similar publications contain many articles that have no more direct relevance to gene therapy than has any study of viral infection or gene expression. This reflects the stage of the field, where any research on experimental exogenous gene expression, or viral transduction, can be 'defined' as gene therapy. The editors and publishers clearly hope the field will take off but don't quite know how or when.

The Journal of Gene Medicine claims to be cross-disciplinary, and includes interesting interviews with prominent scientists in the field, industry updates covering news from gene-therapy companies, and lists of relevant patents. Additional information is available from the associated Internet resources, which include an online database of worldwide clinical trials in progress and free preprint access to all articles that have been accepted and are in press.

The new American counterpart, *Molecular Therapy*, is similar but strong on



methods, which it publishes as separate articles. It also includes up-to-date news and commentaries, with less emphasis on regulatory issues than *Human Gene Therapy*. The research articles are mostly on viral gene therapy with clinical applications, largely from American laboratories. For *Molecular Therapy* to prosper in the long term it may consider broadening its scope and including special features, something *The Journal of Gene Medicine* has accomplished with its Internet resources and journal format.

Clearly, the two established journals cannot rely on age and reputation alone. And the new journals will have the loyalty of their society members, who also provide a ready-made subscription list that will attract advertisers. *The Journal of Gene Medicine* is attractive because its speed of publication is around two to three weeks from acceptance — as against three months for *Human Gene Therapy* and *Molecular Therapy*, and five months for *Gene Therapy*.

All four journals have good review articles, which are the most interesting pieces in many issues, as the significant or 'break-through' papers are published in high impact-factor journals such as *Cell*, *Nature Genetics*, *Science* or *Nature Medicine*.

Will there be enough research in gene therapy to justify the existence of four journals rather than two? The high degree of biotech interest in the field should sustain a lot of hope (and growth) during the coming decade. Let us hope it is justified by clinical usefulness some day.

Genetics in Medicine, a new clinical genetics journal sponsored by the American College of Medical Genetics, is in a different category, as each issue contains only a few articles, which are highly specialized and many quite long. The journal has not yet found its direction. The editors have not matched the crispness of the *Journal of Medical Genetics*, one of its two main existing rivals, or the scientific depth of the more clinical articles in *The American Journal of Human Genetics*. The interesting commentaries on ethics and communication are not even featured on the cover! There may well be a need for another

medical genetics journal, but the editors should define its role. One less attractive feature of the journal is its unusual \$30 per page cost to authors.

♦ <http://www.apnet.com/moltherapy>

♦ <http://www.wiley.co.uk/genmed>

♦ <http://www.com/GIM>

Karin Sitte and Robert Williamson are at the Murdoch Childrens Research Institute, Royal Children's Hospital, Parkville, Victoria 3052, Australia.

Reborn for a new era

Genesis: The Journal of Genetics and Development

editors Richard Behringer & Terry Magnuson

Wiley. 12/yr. Institutional: \$1,395 (US), \$1,617 (outside N. America); new institutional subscribers, \$995 & \$1,217, respectively; individual: \$60 (US), \$132 (outside N. America), to those whose libraries subscribe

Rudi Balling

"Genesis" — What a name! The new journal was 'reborn' in January this year, as volume 26, number 1, from *Developmental Genetics*. According to the editors, *Genesis* wants to break away from the past and leap forward into the functional-genomics era.

The chances of success in this are good. Many of the exciting breakthroughs forming the toolbox of functional genomics come from the labs of developmental biologists: transgenics, chimaeras, stem cells, chemical mutagenesis.

Genetics was extremely powerful in dissecting developmental processes. But, as the editors Richard Behringer and Terry Magnuson say: "Understanding the function of genes and the roles they play in complex biological processes needs more than a study of primary model organisms". Indeed, the first



issues live up to this assessment: in addition to articles on mice, fruitflies, zebrafish and the roundworm *Caenorhabditis elegans*, we find interesting papers on amphioxus, pigs, turtles and the marine ciliate *Euplotes crassus*.

Of course, there are other journals that cover developmental genetics, such as *Development*, *Mechanisms of Development* or *Development, Genes and Evolution*. However, *Genesis* seems well positioned in this journalistic ecosystem. An interesting feature is the "Technology Reports". These cover topics such as the latest tricks for gene targeting in mice or gene silencing in flies, as well as explaining the relevance of the Swiss-Prot database to developmental research — from the sequence of nucleotides to the sequence of pathways.

Genesis seems to have the mix of ingredients required for survival among journals of functional genomics: dedicated editors, rapid publication and an anticipation of the direction of an emerging field. I am sure I will revisit many of the articles I read — some of them to help my students design experiments, some to help me take advantage of the power of the comparative approach.

The editors encourage researchers studying "non-traditional" organisms to submit their papers to *Genesis*. I encourage those with a curiosity in the latest developments in functional genomics to look into the journal.

♦ <http://www.interscience.wiley.com/jpages/1526-954X>

Rudi Balling is at the Institut für Säugetiergenetik, GSF – Forschungszentrum für Umwelt und Gesundheit, Ingolstädter Landstrasse 1, 85758 Neuherberg, Germany.

Psychic secretions

The International Journal of Neuropsychopharmacology

editor-in-chief B. Lerer

Cambridge University Press. 4/yr. £112 (institutional), £56 (individual)

John C. Marshall

It was a slow journey from the Hippocratic humours (black bile, phlegm, blood and yellow bile) to Otto Loewi's 1924 discovery that synaptic transmission from the vagus nerve to the heart muscle was chemically controlled — the first of many neurotransmitter substances (acetylcholine) had been found.

A slightly longer and more tortuous journey led to the conjecture that too much or too little of various neurotransmitters and other neuroactive chemicals was implicated in many neurological and psychiatric dis-

eases. This was reinforced by the development of animal models of these diseases. The upshot was the emergence of a new discipline — neuropsychopharmacology — that gave substance to the claims of Pierre-Jean-Georges Cabanis, an eighteenth-century physician and ideologue, who had notoriously argued that the brain digests impressions and secretes thought.

The *International Journal of Neuropsychopharmacology* covers the entire range of topics that plausibly fall within the domain of its title. A fairly rapid publication schedule combined with high-quality papers make this journal essential reading for all basic and clinical scientists concerned with the biochemistry of the nervous system in health and disease. Thus far, the largest number of research papers in the journal deal with the biochemistry of major depression and schizophrenia. But the wider scope of neuropsychopharmacology is reflected in contributions on movement disorders, mania, anxiety, aggression and pain.

Discussion is facilitated by excellent review articles and a timely trends and perspectives section. The latter feature has included a fine exposition of the health implications of cannabis consumption. Relevant aspects of molecular genetics are also well covered, including a running review of the uses of knockout mice (although neuroscientists who are puzzled by how studies of rats can illuminate schizophrenia will doubtless remain so).

Watch this space. As the editorial team write in their introduction to the journal: "tenets held basic to contemporary neuropsychopharmacology could turn out to be substantially overemphasized, unacceptably simplistic or even incorrect, in the relatively near future."

♦ http://www.journals.cup.org/owa_dba/owa/ISSUES_IN_JOURNAL?JID=PNP

John C. Marshall is at the Neuropsychology Unit, University Department of Clinical Neurology, Radcliffe Infirmary, Oxford OX2 6HE, UK.

Problems of a glandular nature

Pituitary

editor-in-chief Shlomo Melmed

Kluwer. 4/yr. \$314 (institutional), \$100 (individual)

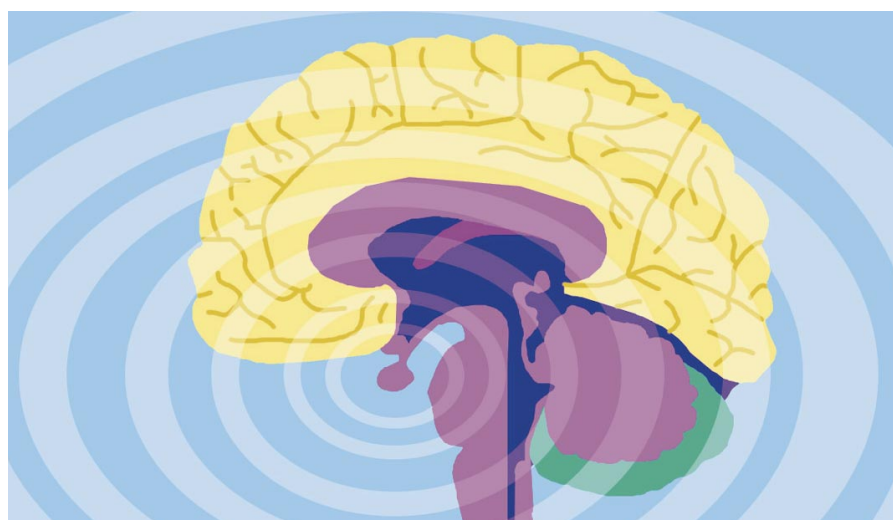
Stafford L. Lightman

The editorial board includes some of the best-known international figures working on the pituitary gland — clinical endocrinologists, neurosurgeons, molecular biologists, neuroendocrinologists and pituitary histopathologists. *Pituitary* is aimed at all areas of hypothalamic-pituitary function, from the most fundamental cell biology through to clinical case reports.

The journal must have given its editors considerable problems. There are already many journals covering clinical endocrinology, neuroendocrinology and neurosurgery, so it was important for them to carve out a niche that would serve a population of clinical and basic scientists. They have sought to do this by dividing the journal into sections devoted to basic research, clinical studies and reviews, with occasional 'portraits' of recent congresses or other newsworthy events.

Pituitary itself is nicely presented, with good quality print, although some of the plates of immuno-stains lack clarity. The editors have clearly had a struggle to attract the high-quality manuscripts the journal deserves, and this may reflect the catch-22 situation of awaiting cover by the citation and abstracting agencies.

There are interesting papers at the clinico-pathological interface, which could be the journal's greatest strength. It is, however, noticeable that although the initial volumes had quite a large basic science content, this seems to have diminished in later issues, and although the journal is published only four times a year, the most recent volume in May



book reviews

2000 is only 50 pages long, has two basic science papers and five clinical articles, of which three are case reports.

The scientific and clinical excellence of the editors puts them in a good position to attract the high-quality manuscripts essential to making this journal a success. At present they have their work cut out for them.

► <http://www.wkap.nl/journalhome.htm/1386-341X>

Stafford L. Lightman is in the Department of Medicine, University of Bristol, Bristol Royal Infirmary, Bristol BS2 8HW, UK.

The art of the saveable

Animal Conservation

editors Michael W. Bruford,
John L. Gittleman, Georgina M. Mace
& Robert K. Wayne

Cambridge University Press. 4/yr. £89, \$145
(institutional), £41, \$69 (individual),
£19, \$31 (student)

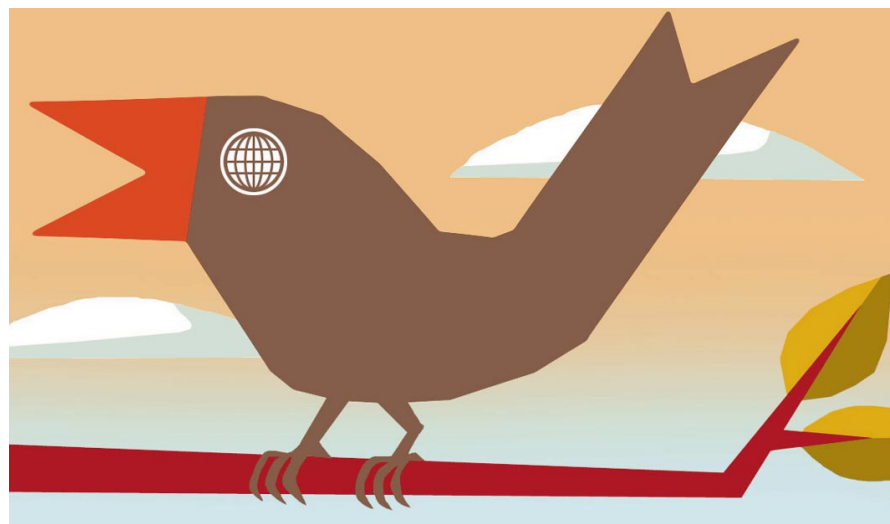
William J. Sutherland

If politics is the art of the possible and science the art of the soluble, then conservation is surely the art of the saveable. The widespread concern over local and global environmental change has produced an outburst of interest among students, a redirection of ecological research towards applied questions and an expanding community of conservation biologists.

Animal Conservation is a very welcome contribution to the art. It covers the spectrum of conservation science (with a particular strength in evolution and genetics) and balances practicalities with fundamental issues. Thus the latest issue includes the genetic management of chondrodystrophy in California condors, whether fluctuating asymmetry can be used to detect inbreeding in endangered species, a protocol for re-establishing the beaver in Scotland and using human densities to interpret declines of large carnivores. The contents are usually so interesting and topical that I find myself reading most of the journal, and the students on our MSc course in conservation biology successfully petitioned our librarian to start subscribing!

Animal Conservation focuses on conservation science and excludes associated subjects such as policy and education. It is thus more narrow in scope than the market leader, *Conservation Biology*. The restriction to animals seems somewhat old-fashioned, especially as so many concepts apply equally to the rest of biodiversity; general papers on communities, biodiversity and ecosystem processes are likely to be submitted elsewhere.

The journal has started well. Even the first



issue contained a series of interesting papers, and this quality has been maintained. *Animal Conservation* has the potential to become the major conservation science journal within a few years.

► <http://uk.cambridge.org/journals/ani>

William J. Sutherland is at the School of Biological Sciences, University of East Anglia, Norwich NR4 7TJ, UK.

Strong growth in the environmental field

Journal of Environmental Monitoring

editor Harp Minhas

Royal Society of Chemistry. 6/yr. £350, \$578
(institutional), £60, \$96 (individual)

Journal of Environmental Policy & Planning

editors Kevin Bishop, Andrew Flynn
& Terry Marsden

Wiley. 4/yr. \$190 (institutional), \$105
(individual)

Graham Wood

Environmental monitoring and the implementation of informed and effective environmental policies are essential components of any initiative that attempts to grasp the illusive Holy Grail of sustainable development. Continued growth in the status of environmental issues seems unquestionable, and a huge diversity of research activity now exists that could come under the environmental umbrella. The inherent interdisciplinary nature of environmental research means that material in an emerging sub-discipline has a tendency to be scattered among a plethora of journals. Reflecting the natural-science and social-science traditions respectively, the *Journal of Environmental Monitoring* (JEM) and the

Journal of Environmental Policy & Planning (JEPP) both seek to provide a single coherent outlet for research within their individual evolving fields.

Published by the Royal Society of Chemistry, JEM focuses on all aspects of the measurement of chemical, physical and biological agents indoors, outdoors and in the workplace. The emphasis of the journal is on exposure assessment in relation to adverse environmental and health effects. To date, the journal has published contributions from a diverse international group of writers, featuring a good-quality mix of long and short research articles, many derived from programmes funded by national research councils, the European Commission and major industrial sponsors. Production quality is high and the journal can boast an impressive turnaround time, with an average duration between submission and publication of just four months. In addition to research papers, JEM incorporates a variety of short, magazine-style scientific articles covering topical issues of general interest, and a useful news section that often includes web links to further information. All in all, JEM is a well-rounded periodical that has got off to a good start in its first year.

JEPP aims to provide "a forum for critical analysis" in the field, a key theme being to facilitate integration of different communities (including academics and policy-makers), different policy sectors and spatial scales, from local through to global. The journal is primarily, though not exclusively, concerned with Europe and has attracted a range of good-quality articles. It covers a variety of theoretical and empirical studies. A useful feature is the "progress report" in each issue, which has either a thematic or a regional focus. It is a healthy sign that the editors have not strayed from the originally stated aims and scope of the journal, and although submission-to-acceptance times are painfully slow, they are not atypical (five to 18 months, with a further publication lag

of two to four months). At \$190 the journal represents good value, and it is pleasing to see that is moving from three to four issues in the forthcoming year. Overall, *JEPP* would certainly be a worthwhile addition to library stocks for institutions with specific research or teaching interests in environmental policy. ■

♦ <http://www.rsc.org/is/journals/current/jem/jempub.htm>

♦ <http://www.interscience.wiley.com/jpages/1523-908X>

Graham Wood is in the Impacts Assessment Unit, School of Planning, Oxford Brookes University, Oxford OX3 0BP, UK.

More than a breath of fresh air

Wind Energy

editor-in-chief Robert W. Thresher
Wiley. 4/yr. \$320 (institutional),
\$245 (individual)

Ian Fells

The wind bloweth where it listeth; and thou hearest the sound thereof, but canst not tell whence it cometh and whither it goeth. St John, ch. 3, v. 8

Man has harnessed the wind to power sailing ships, turn windmills to grind corn and, latterly, to generate electricity. Wind is fickle and unpredictable, as those who race yachts may know to their cost. But now wind energy, or rather, electricity generated by huge wind turbines — each producing two megawatts of power — is big business and growing at 30% a year. Worldwide there are 15 gigawatts of wind-generated electricity capacity.

As is so often the case, the practical engineering is well ahead of the science. The international journal *Wind Energy* has been set up to try to redress this imbalance and “offers a major forum for the reporting of advances in this rapidly developing technology ... to



harness clean energy from the wind”.

The papers range widely, from “Wind power meteorology”, through “Modelling methods for wind turbine wakes” to “Mature offshore wind technology”. They are generally rigorous, with a nice balance of theoretical and practical content. The fluid mechanics of wind systems is crucial for energy to be generated efficiently and the journal’s contents reflect this. Surprisingly, despite the journal’s mission statement, the economics of wind-generated electricity and socio-political issues seem not to feature as yet, although a section on “Broader perspectives” is promised.

There is some competition from *Wind Directions*, a journal published by the European Wind Energy Association, and from *Renewable Energy World*, published by Edward Milford, but these are essentially trade magazines and so *Wind Energy* fills an important gap. ■

♦ <http://www3.interscience.wiley.com/cgi-bin/jtoc?ID=6276>

Ian Fells is principal consultant at Fells Associates, 29 Rectory Terrace, Newcastle upon Tyne NE3 1YB, UK.

Chemists exploring a verdant field

Green Chemistry

editorial board chairman Roger A. Sheldon
Royal Society of Chemistry. 6/yr. £250, \$412
(institutional, including site-wide electronic access), £45, \$72 (individual)

Walter Leitner

“Green chemistry” may sound like an oxymoron to those who identify chemistry with cheap and nasty industry, and associate “green” with fundamental activists fighting against technological progress. But the need for clean and environmentally benign processes based on a sustainable development is now widely recognized and has found its way into many corporate targets and research policies. In this context, “green chemistry” is defined as the utilization of a set of principles that reduces or eliminates the use or generation of hazardous substances in the design, manufacture and application of chemical products (P. Anastas & J. C. Warner, *Green Chemistry*, Oxford Univ. Press, 1998).

Despite its seemingly clear focus on technological application, the above definition results in a tremendous challenge for fundamental science in chemistry. What impact do chemicals have on the environment? How can we develop alternatives to potentially hazardous reagents? Is it possible to replace traditional reaction sequences resulting in the inevitable generation of by-products with new transformations where all starting materials are incorporated in the product? Can we design, or efficiently screen for, new catalysts that make reactions cleaner or help to save energy? What are the chances of replacing organic solvents with more benign reaction media? These are just some of the questions that organic and inorganic synthesis, physical chemistry and chemical engineering can help to address in this field.



book reviews

The new journal *Green Chemistry* provides an international forum for discussion and scientific exchange of this fairly heterogeneous community, which is held together by a common goal rather than a specific methodology or a specialized subject.

Each of the six annual issues contains two sections—News & Views and Research Publications. The first, slightly shorter, part highlights new developments in technology, research policy or science and provides information about symposia, funding programmes and ongoing projects. In the second part, about ten scientific papers deal mainly with the topics described above and include experimental procedures, usually in detail. The quality of layout and graphics is excellent throughout the journal, making it easy to scan the contents if time is short. The reasonably priced institutional subscription includes site-wide electronic access, and additional personal subscriptions to the printed version are available at very moderate rates.

Owing to its rather broad definition, the field of green chemistry is still in flux: in this situation, the publication of a periodical plays an important role in delineating the exact scope of a field. At the moment, each research article in *Green Chemistry* is accompanied by a box in which one of the editors describes the “green context” of the work. The journal is well on the way towards making the need for this explanation obsolete, and has the potential to establish its position in the chemical literature. ■

♦ <http://www.rsc.org/ls/journals/current/green/greenpub.htm>

Walter Leitner is at the Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, D-45470 Mülheim an der Ruhr, Germany.

Chemistry's voice heard at last

Foundations of Chemistry: Philosophical, Historical, Educational and Interdisciplinary Studies of Chemistry

editor-in-chief Eric R. Scerri
Kluwer. 3/yr. \$124 (institutional),
\$70 (individual)

Bernadette Bensaude-Vincent

Twentieth-century philosophy of science—especially the anglophone tradition—virtually ignored chemistry. From logical positivism through to Popper, Kuhn, Lakatos and Feyerabend, physics has been the model science. This neglect of the philosophy of chemistry prompted a strong reaction in the 1990s in Britain, Germany and the United States. After a number of conferences and symposia in 1994, two journals were created

to fill the gap: *Hyle: An International Journal for the Philosophy of Chemistry* in 1995, published by the University of Karlsruhe in Germany, and *Foundations of Chemistry* in 1999.

If we admit that ‘filling a gap’ is more than a rhetorical justification for new journals, the main question is, ‘how do you fill the gap?’. The answer is not clear from the four issues of *Foundations of Chemistry*. Its editor-in-chief Eric Scerri, an American chemist, turned to the philosophy of chemistry in order to fight against the reduction of chemistry to quantum physics. His initial motivation was to demonstrate that the foundations of chemistry lie in chemistry itself. For him, the philosophy of chemistry is mainly a strategy to claim chemistry's autonomy.

The journal's title clearly indicates that it is a response of the chemical community to the longstanding neglect of chemistry which began with *Foundations of the Unity of Science*, the famous publication of the logical positivism movement. Scerri, aware of the dangers of a reactive attitude, is trying to broaden the scope of the journal to include such topics as the role of instruments in chemistry, and to develop a historical dimension lacking in logical positivism. However, the success of the enterprise will rely on the journal's ability to interest philosophers and historians of science. Its future depends on the formation of a multidisciplinary and multicultural research community of philosophers of chemistry. ■

♦ <http://www.wkap.nl/journalhome.htm/1386-4238>

Bernadette Bensaude-Vincent is in the Département de Philosophie, Université Paris X-Nanterre, 92001 Nanterre Cedex, France.

Struggling with complexity

Crystal Engineering

editor-in-chief Mike Zaworotko
Pergamon-Elsevier. 4/yr.
euro 166.54, \$186

Gastone Gilli

Reductionism has been very successful when applied in a descending order of scale and complexity, the normal direction for reducing biology to chemistry, solid-state physics and, finally, elementary-particle physics. Unfortunately, as remarked in 1972 by physics nobellist Philip Anderson, “the ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and



reconstruct the universe ... Instead, at each level of complexity, entirely new properties appear.” This makes it next to impossible to deduce the complexity and novelty that can emerge through knowing the composition of many elementary entities.

Anderson's views are regularly verified in the new cross-disciplinary field of crystal engineering. The related journal is aimed at developing protocols for predicting and controlling the structure and function of molecular crystalline materials. Although the basic laws governing molecular interactions at the sub-microscopic level are essentially known, the preferred forms of molecular aggregation at the upper level of complexity (the macroscopic crystal) remain substantially unpredictable.

This may justify the journal's editorial policy, which focuses on the synthesis and determination of the structure of molecular crystals, co-crystals and clathrates that have specific and potentially useful internal patterns. There is less emphasis on theoretical interpretation or *ab initio* crystal structure prediction. The journal therefore adds to the harvest of experimental data from which the final rules of crystal composition will, hopefully, emerge.

Crystal Engineering has other merits, in particular a pleasing typographic form and a truly outstanding list of editorial members. Its success will depend on its ability to complement other well-established crystallographic journals by becoming a reference for general problems of crystal design, such as the systematics of molecular interactions, the topological features of crystals that give them their specific physical properties, and the rationalization of the often cryptic supramolecular nomenclature. ■

♦ <http://www.crystal-engineering.net>

Gastone Gilli is in the Department of Chemistry, Centre for Structural Diffractometry, University of Ferrara, Via Borsari 46, I-44100 Ferrara, Italy.



Across the boundaries

Interfaces and Free Boundaries: Modeling, Analysis and Computation

editor Mario Primicerio
Oxford University Press. 4/yr. £195, \$335 (institutional), £176, \$302 (institutional, online only), £50, \$80 (individual)

Albert-László Barabási

The alarming loss of biodiversity is accompanied by an increasing diversity of scholarly thought, as demonstrated by the rate at which specialized journals emerge. The survival of free thinking is guaranteed by safe havens for those sharing the same interests, tools or ideas, but they are scattered across disciplines.

Interfaces and Free Boundaries is one such safe haven, offering a fertile soil for applied and pure mathematicians trying to reach a common language in this rather fragmented and interdisciplinary field. The apparently eclectic selections of contributions, ranging from superconductivity to dendritic growth, offers a culture shock at first. However, the source of this diversity is hidden in the subtitle: *Modeling, Analysis and Computation*. And indeed, the journal does offer common ground for those interested in the computational and mathematical tools for dealing with interfaces and free boundaries, irrespective of discipline. As a result, experimentalists should beware: the papers clearly favour rigorous mathematics above phenomenology — probably in line with the intent of the editorial board, the great majority of whom come from applied maths departments.

But should you be an experimentalist or theorist faced with a free-boundary problem, this is the place to look for expertise, papers and ideas — the four issues published so far offer a wide spectrum of articles that touch on most of the techniques needed to master this field. And all this in a single journal, doing away with the need for often fruitless database searches or trips to various libraries across the campus. The best way to get the flavour of the journal is to visit its online edition, which displays the table of contents of the issues published so far. In this respect, Oxford University Press has done an excellent job, giving the journal all the perks it needs for survival, including an online edition, electronic submissions and a pleasing appearance, with the possibility of colour pictures.

Should you run to your librarian for a subscription? Look at the online version before you drop any other journal from the shelf. But if your budget can be stretched to include it, it will be an unselfish decision for

any department, since people from other departments might wish to peek too. ■

♦ <http://www3.oup.co.uk/infree/?LO>

Albert-László Barabási is in the Department of Physics, University of Notre Dame, Notre Dame, Indiana 46556, USA.

Europe's hard forum for soft matter

The European Physical Journal E — Soft Matter

editors-in-chief A. M. Donald, J.-F. Joanny, M. Möller & G. Reiter
Springer. 12/yr. \$1,181 (US), DM1,872 plus carriage charges (Germany DM37.20; elsewhere DM52.80)

Steve Granick and Sung Chul Bae

The study of 'soft matter' has burgeoned as an academic discipline during the past generation. Its prominence has matched the study of 'hard matter' — traditional condensed-matter physics. The scope of 'soft matter' is larger than physics, however. The preoccupation is with systems (polymers, biopolymers, membranes, self-assembled nanostructures) where the energetic interactions are commonly so 'soft' that complexity, disorder and the capacity to manipulate and tune structure and dynamics emerge in fundamentally new ways. The worldwide community is struggling to deal with the fact that these systems straddle so large a spectrum of traditional disciplines: not just physics, but also chemistry, biophysics and some mechanics.

Seven years ago these considerations led the American Physical Society to launch a new journal, *Physical Review E*, which immediately established itself as a premier archival journal in terms of quality and prestige. It is altogether fitting that the European Physical Society should do the same.

Here three distinguished European journals — *Il Nuovo Cimento D*, *Journal de Physique* and *Zeitschrift für Physik B* — seek rebirth in the launch of a blue-ribbon European journal devoted to this new discipline.

The journal's ambitious goal is to develop a wider readership and more eclectic submissions than would be expected of a physics journal. Among the articles published during the initial six months the quality is high, but the flavour is not yet really different from that of *Physical Review E*. Representation from other subfields of soft matter will probably grow. Every scientific research library will want to hold this blue-ribbon archival journal. ■

♦ <http://www.link.springer.de/journals/epje>

Steve Granick is in the Departments of Materials Science and Engineering, Physics, Chemistry and Chemical Engineering, University of Illinois. Sung Chul Bae is in the Department of Materials Science and Engineering, University of Illinois, Urbana, Illinois 61801, USA.

Making way for complex issues

Nonlinear Phenomena in Complex Systems

editors-in-chief V. A. Gaisyonok & F. I. Kuvshinov

Education & Upbringing, Belarus. 3/4/yr. \$158 (institutional), \$98 (individual & institutions from developing countries)

William Ditto

Small changes can often have fundamental effects on the behaviour of complex systems. This is self-evident in the proliferation of interdisciplinary journals devoted to nonlinear dynamics and complex systems. Amid this creative renaissance (or journalistic excess, depending on your viewpoint), new journals



book reviews

must provide a service to the scientific communities rather than just add to the overall explosive numbers of scientific papers.

Nonlinear Phenomena in Complex Systems delivers to the community in a couple of ways. First, it provides a mostly regional forum (primarily Central and Eastern European) for both experimental and theoretical research in interdisciplinary complex systems. Thus it gives both an editorial and a scientific voice to the many creative individuals who were previously relegated to the cloistered world of Soviet science journals. Lest those of us in the United States get too uppity, we should remember that it was only a few decades ago that *Physical Review* was considered a regional or 'colonial' journal.

But the journal must ultimately stand on quality alone, and the quality and content of the papers in *Nonlinear Phenomena* are rather uneven. One problem lies in the presentation of technical data. Topics covered range from esoteric mathematical proofs and formalisms to practical experimental research. Unfortunately, there seems to have been little editorial attempt to persuade the authors to make their papers more accessible to an interdisciplinary audience. This is an extremely important aspect and very hard to pull off.

Apart from issues of content, there are also problems with the quality of presentation. Specifically, the figures range in quality from adequate to unreadable. This at least must be addressed for the journal to survive.

On the positive side, there are articles from regional conferences that are quite stimulating. Clearly, a reader interested in the latest in interdisciplinary complex systems research would do better initially to peruse this journal's competitors, for example *Chaos* and the *International Journal of Bifurcations and Chaos*. But *Nonlinear Phenomena* is an improving publication that serves a regional audience with papers that range from the banal to the fascinating. ■

► <http://ccisc.bas-net.by/npcs>

William Ditto is in the Georgia Tech/Emory Department of Biomedical Engineering, 315 First Drive, Atlanta, Georgia 30332-0535, USA.

Backing the n-word

Journal of Nanoparticle Research

editor-in-chief Mihail C. Roco
Kluwer, 4/yr. \$278 (institutional),
\$88 (individual)

Saul Tendler

When Shakespeare penned the line "I do not like her name", he could have been thinking about nanotechnology. Derived from the Greek *nanos* for dwarf, it may revolutionize our lives through the creation of nanometre-scale sensors, devices and machines. How-

ever, early claims about the potential of the subject leave many people cautious about using the n-word.

The *Journal of Nanoparticle Research* seeks to embrace all things nanotechnology while avoiding its name. With a hard focus on nanoparticles, the scope of the journal includes fundamental science, chemical synthesis, modelling, simulations, instrument development and molecular/particulate self-assembly into higher-order architectures. This broad remit extends from the biological and chemical to the physical sciences. The first volume of the journal has attracted an impressive number of full papers, many from leading groups with international reputations. Perhaps unsurprisingly, the majority of the papers cover the production and properties of inorganic nanoparticle systems.

The journal carries both long and short papers, reviews, technology notes and meeting reports. There are also policy papers on funding initiatives. Intriguingly, the journal allows papers directed towards education: the first edition has a splendid paper on making the nanoworld comprehensible for schools and beyond. Given the publication's specialist nature, though, it may be that some of this outreach activity will fall away over time.

There are many top-quality physico-chemical journals that would publish much of the work found in the first volume. But such publications do not attract a sufficient critical mass of papers in this area to create a comprehensive view. With quality basic science, and applications assured in areas as diverse as catalysis and drug delivery, *Journal of Nanoparticle Research* provides a cross-disciplinary focal point for the subject. Success is assured if it can maintain this focus. ■

► <http://www.wkap.nl/journalhome.htm/1388-0764>

Saul Tendler is at the School of Pharmaceutical Sciences, University of Nottingham, Nottingham NG7 2RD, UK.

Plant power against pollution

International Journal of Phytoremediation

editor-in-chief Guy R. Lanza
CRC Press, 4/yr. \$385 (institutional, print plus online), \$295 (institutional, print or online), \$125 (to members of the Association for the Environmental Health of Soils)

William Purvis

Phytoremediation relies on the natural properties of plants to help clean up hazardous metal and organic wastes. Bioremediation with microorganisms was used in the Exxon Valdez oil spill in Alaska and also after the Chernobyl nuclear power plant accident. The idea of using plants to extract, inactivate, transform or degrade contaminants is appealing — they bring 400 million years of evolutionary processes across hugely diverse groups. Moreover, increased environmental awareness and the need for sustainable development have sparked intense activity in this area across the world, particularly in the United States.

Appropriately, the first issue of *International Journal of Phytoremediation* features an invited and readable review article by renowned ecotoxicologist André Sobolewski, who discusses the biological and non-biological processes responsible for removing metal from mine drainage by wetlands. This is the classic application, and uptake by plants is only part of the story. Bacteria may accumulate and transform metals which may also be sorbed to organic matter and inorganic materials.

Phytoremediation is an interdisciplinary subject, its success depending on diverse areas of science that include genetics, physiology, biochemistry, microbiology, soil science, geochemistry and engineering. This journal necessarily takes a broad view



and encourages the submission of papers describing technologies that combine chemical and other processes with phytoremediation.

The role of the microorganisms involved is well treated and geochemistry has not been neglected, trends that I hope will continue. Ways of enhancing metal availability to improve its uptake by plants are also discussed. The journal has achieved a sensible balance between original research articles, invited reviews, special commentaries and technical notes. The fact that much activity is currently US-based may explain the strong preponderance of US scientists in the impressive editorial board, although leading scientists from other countries are also represented.

This journal is the first of its kind devoted to phytoremediation. Its appearance is timely considering the importance of the subject and the rapidly growing literature at present scattered across journals devoted to diverse disciplines. It should allow improved communication between scientists and lead to the further collaborations necessary to extend our knowledge in this field. The journal deserves to succeed and will undoubtedly be of value to many life and Earth scientists.

♦ <http://www.crcpress.com:80/us/jour/jourinfo/15126514.asp?mcsid=>

William Purvis is in the Department of Botany, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

Green and red lights at the cell

Traffic: The International Journal of Intracellular Transport

editors Frances M. Brodsky, Mark C. P. Marsh & Sandra L. Schmid
Munksgaard. 12/yr. \$380 (institutional), \$120 (individual)

Reinhard Jahn

Since the classic work of George Palade and colleagues in the sixties, the field of cell membrane traffic has undergone an almost exponential growth. Milestones include the introduction of yeast genetics for identifying genes involved in individual trafficking steps and the development of cell-free assays for investigating membrane transport. More recently, the dramatic progress in live-cell microscopy and the inroads of structural biology have transformed our understanding of trafficking proteins.

Traffic is the first journal devoted entirely to the coverage of membrane traffic. Launched at the start of this year by a group of renowned scientists including Frances

Brodsky, Mark Marsh, Sandy Schmid and the late Thomas Kreis, the goal is to create "a central journal to gather together publications that are of most interest to those working on intracellular trafficking".

Published monthly, about half of the journal space is devoted to short and well-illustrated reviews centred around a common theme. Other features include a "toolbox" section covering technical developments (recently including galleries of crystal structures of trafficking proteins, and a collection of John Heuser's stunning three-dimensional electron micrographs), meeting reports, and commentaries by senior scientists about current issues of general interest. The remainder contains research papers.

Where does *Traffic* stand after the first seven issues? Obviously, it is very difficult to start a new journal in a highly competitive field. So far, *Traffic* publishes only three to five research papers a month, too few to have a serious impact. In a world driven by 'impact factor' mania, students and postdocs are not easily persuaded to submit high-quality papers to a new journal with a still uncertain acceptance and will turn instead to established cell biology journals. However, it is refreshing to see senior scientists serving the community as enthusiastic editors in a journal market increasingly dominated by professional editors with limited research experience and by revenue-oriented business managers.

Despite its (still) small size, *Traffic* has succeeded in attracting many leading scientists, as can be seen by a glance at the editorial board and, more importantly, by the list of contributors.

The journal's appeal is heightened by an attractive layout and excellently reproduced drawings and images, mostly in colour. *Traffic* still has a long way to go before it becomes a 'must read' for the trafficking community, but it has made a strong start and is well on its way. All it needs now is its growing accep-

tance as an attractive forum for the publication of excellent research articles.

♦ <http://journals.munksgaard.dk/traffic.nst/alt/forside-1304>

Reinhard Jahn is in the Department of Neurobiology, Max-Planck-Institute for Biophysical Chemistry, 37077 Göttingen, Germany.

The nidus and the crucible

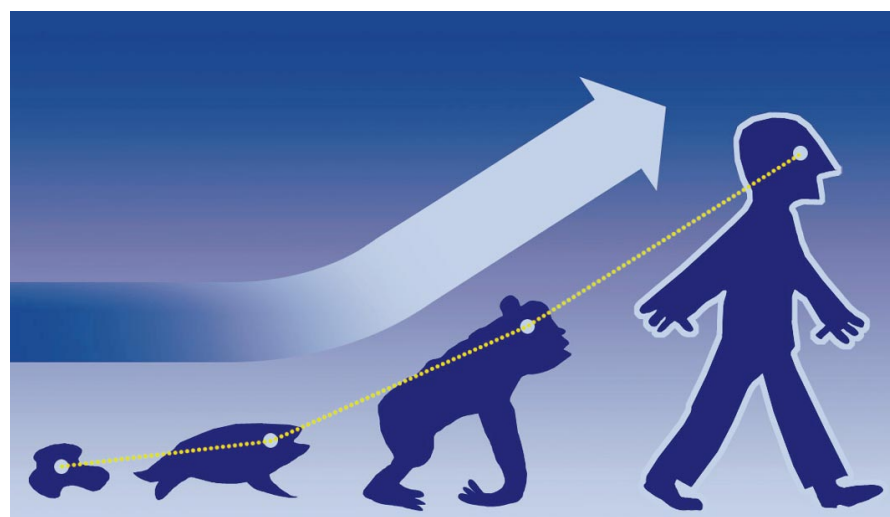
Evolution & Development

editor-in-chief Rudolf A. Raff
Blackwell Science. 6/yr. Print only:
institutional: \$195 (US), \$215 (elsewhere);
individual: \$80 (US), \$100 (elsewhere).
Combined print/online: institutional:
\$214.50 (US), \$234.50 (elsewhere);
individual: \$88 (US), \$108 (elsewhere)

Moya Meredith Smith

This journal emerges from the melting pot of our knowledge of how genes act during development to produce such a diversity of characteristics. It is an exciting new arrival on the scene. At the cutting edge of integrated biological science, it offers us a merging of two disciplines separated by a century of dispute. Will it bring the 'reciprocal illumination' long advocated by the classic theory that an understanding of development will illuminate that of evolution, and vice versa?

There is increasing interaction between the separate disciplines of palaeontology, population biology, developmental biology and molecular biology. And the journal aims to show how evolution and development are intimately related despite being fundamentally different processes. However, these separate disciplines have distinct and opposite concepts. One emphasizes diversity and the elimination of genetic options by natural



book reviews

selection to produce a distinct animal group. The other finds universally similar processes with similar genes acting to produce an individual animal.

Is it, as the journal's first editorial said, "the birth of a discipline"? Perhaps these two partners are too disparate to provide a unified science. But this attempt to create an intellectual home for the "new synthesis" will make fascinating reading. All the key players are there in the first five issues of the journal. The format is a couple of forum papers and a review, with the rest being mostly original research papers.

Examples of issues raised range from the ancient origins of axial patterning genes through to patterns and rates of vertebral evolution in amphibians and the evolution of the embryonic dorso-ventral axis. Although there is a diverse mix of contributions, spread evenly between the four disciplines, the consistent character is a view based on evolutionary perspectives.

The constraints on the future of this journal will be those of treading a tightrope in the arena of development that will produce an integrated science of biological form. The risk is descent into gene function at the molecular level. The gain would be to achieve the birth of a new platform from which to explain why fossils, with their unimaginable diversity of form, are not just isolated examples of failed experiments in the crucible of life. ■

♦ <http://www.blackwellscience.com/journals/evolution/index.html>

Moya Meredith Smith is in the Department of Craniofacial Development, Dental Institute, King's College London, University of London, London SE1 9RT, UK.

Crossing the resistance divide

Drug Resistance Updates

editor-in-chief Nafsika Georgopadakou
Churchill Livingstone. 6/yr. \$225, £180
(institutional), \$115, £85 (individual)

Eric J. Arts

Although resistance is a common and growing concern in the treatment of infectious disease and cancer, *Drug Resistance Updates* is one of the first journals to address chemotherapeutic resistance as a main theme and to review a wide variety of resistance-related issues.

The bimonthly issues cover a range of topics and pose related questions. For example, what is the mode of resistance to the antimalarial drug chloroquine and should we take this resistance mechanism into account when prescribing newer drugs such



as mefloquine for malaria prophylaxis or treatment? Is the induction of efflux pumps in azole-treated yeast cells analogous to the overexpression of outer membrane proteins acting as efflux pumps in Gram-negative bacteria or that of the MDR-1 P-glycoprotein in human tumour cells?

Similarities between these mechanisms are often difficult to identify because research on drug resistance has been segregated into various reputable but specialized journals focusing on cancer chemotherapy, microbes or viruses. The editors of *Drug Resistance Updates* have had the good sense to bring together perspectives and reviews on drug resistance in bacteria, cancer, fungi, protozoa and viruses and to provide a simple format for comparing drug-resistance mechanisms between organisms.

The reviews are concise and informative and have recently included descriptive diagrams summarizing complex drug-resistance mechanisms. A good mix of established experts and ambitious newcomers as authors provides a fresh perspective on both old and new drug-resistance topics such as bacterial resistance to β -lactam drugs (for example, penicillin G) and recent evidence for the resistance of influenza virus to neuraminidase inhibitors (such as zanamivir).

Unlike several other review journals,

the areas of microbial and cancer drug resistance are so extensive that there has been little repetition in this journal over its three years of publication. These reasons alone should ensure its survival in a competitive market for high quality and timely reviews.

Although the journal does not solicit comments on previous reviews, it does invite commentaries or perspectives on such topics as managing, characterizing and improving the detection of drug resistance. Even so, more articles comparing drug-resistance mechanisms employed by different organisms would add much to this journal's impact. Such comparisons are invaluable for designing new inhibitors, characterizing resistance mechanisms and developing methods to prolong drug use by reversing resistance. In addition to its content and format, the reasonable subscription price justified my purchase of two subscriptions and the use of this journal as a text in our graduate-level course 'Mechanisms of drug resistance'. ■

♦ <http://www.harcourt-international.com/journals/drug>

Eric J. Arts is in the Division of Infectious Diseases, Department of Medicine, Case Western Reserve University, Cleveland, Ohio 44106, USA.

Also submitted

The following journals were also submitted, but were not reviewed.

- Molecular Cell Biology Research Communications (Academic)
- Physiological Genomics (American Physiological Society)
- Neoplasia (Nature Publishing Group)
- Antioxidants & Redox Signaling (Mary Ann Liebert)
- Marine Biotechnology (Springer)
- Journal of Alzheimer's Disease (IOS Press)
- Journal of Molecular Microbiology & Biotechnology (Horizon)
- Systems Engineering (Wiley)
- Cellular Microbiology (Blackwell)
- Cloning (Mary Ann Liebert)

Down to earth

Why soil — and soil science — matters.

Dan H. Yaalon

Many ancient religions recognized the importance of soils, and their customs evolved into a spiritual attachment to the life-giving Earth. But surprisingly, ancient and classical scholars did not study the nature of soil. Early scientists also ignored it. For instance, the famous naturalist Alexander von Humboldt (1769–1859), a founder of plant geography, never compared the soils of the several continents on which he studied the distribution of plant species. This attitude still crops up frequently. The *Fontana History of the Environmental Sciences* (1992) contains no mention of soils as a branch of environmental science, although other Earth sciences are included. Is soil just dirt, too commonplace for mention or study?

I am a pedologist — an Earth scientist focusing on the origin and distribution of soils in relation to the history of landscapes. We have much to learn about non-arable soils, and must try to integrate our knowledge into a holistic view of the Earth's dynamics and biogeochemical transformations. Soils are economically and socially important. They can even have beauty: the soil scientist Hans Jenny (1899–1992) was enchanted by the soils depicted in paintings. To paraphrase Leonardo da Vinci: why do we know more about distant celestial objects than we do about the ground beneath our feet?

New ideas about the nature and origin of soils emerged only in the second half of the nineteenth century. V. V. Dokuchaev (1846–1903) and E. W. Hilgard (1833–1916), both mineralogists and chemists by training, recognized in their soil surveys that climate, vegetation and substrate were all important, and saw the importance of soil horization — the development of different layers of soil parallel to the surface — in representing and elucidating a landscape's history. Dokuchaev had imperial backing in Russia, and several distinguished followers; Hilgard, although a respected university professor in the United States, was not favoured by the establishment. An opportunity to promote his ideas was lost when J. W. Powell and he failed to establish a geological–agricultural (soil) survey in the US Geological Survey. Language barriers hindered communication between soil scientists, and the spread of knowledge was painfully slow, even after the new Russian pedogenetic ideas were presented at world exhibitions and translated.

Gradually, topographical and biological



Root of life: the soil forms the central link between the climate and biogeochemical systems.

effects, and the duration of soil-formation processes were all recognized as equally important factors in soil evolution. It took more than one generation before C. F. Marbut (1863–1935) included the concepts of external and internal environmental effects on pedogenesis in the influential US Department of Agriculture soil survey, established by Milton Whitney in 1899. When Jenny submitted his now-classic book on the 'five soil-forming factors' and the quantitative approach to single-factor soil-forming functions, it was at first rejected for publication. It took five years before the book was eventually published in 1941.

The importance of soils as a life-support system and in the production of food and fibre was duly recognized. There were spec-

To paraphrase Leonardo da Vinci: why do we know more about distant celestial objects than we do about the ground beneath our feet?

tacular achievements, helping to feed the ever-growing population. Nowadays, most of the 50,000 soil scientists work in agronomic institutes, studying the composition and dynamics of soils in ever greater detail. Yet less than 5% of the global agricultural research budget goes on soil research.

The use of soils in road building, construction, ceramics and the cement and aluminium industries is another area where a basic knowledge of soil and landscapes is important. Technological institutes promote this study, which is anchored in ancient practical applications.

Soils teem with life. The Nobel laureate Selman Waksman (1888–1973) isolated streptomycin from soil biota, and the preservation of pedodiversity and biodiversity may aid similar research in the future. Also, it seems plausible that biological evolution was influenced and constrained by the properties of the soil environment, an attractive field of unexplored research. For Earth scientists, ancient and buried soils are one of the better proxies for reconstructing past climate and the development of the landscape.

But it is as the transformer, regulator, buffer and filter of water, nutrients and other dissolved and dispersed compounds that soils are most important to humankind — a focal and connecting link between the biogeochemical cycles of the Earth and the dynamic atmospheric system. In the conceptual wiring diagram of the International Geosphere–Biosphere Program the soil system, especially its carbon dynamics, is the central link between the physical climate and biogeochemical systems. It is therefore a major route to understanding and predicting the effects of human actions on the Earth. ■

Dan H. Yaalon is at the Institute of Earth Sciences, Hebrew University, Givat Ram Campus, Jerusalem 91904, Israel.

JEREMY BURGESS/SP/L

The Great Goodbye

The emotional cost of the post-evolutionary divide.

Robert Charles Wilson

The hardest part of the Great Goodbye, for me, was knowing I wouldn't see my grandfather again. We had developed that rare thing, a friendship that crossed the line of the post-evolutionary divide, and I loved him very much.

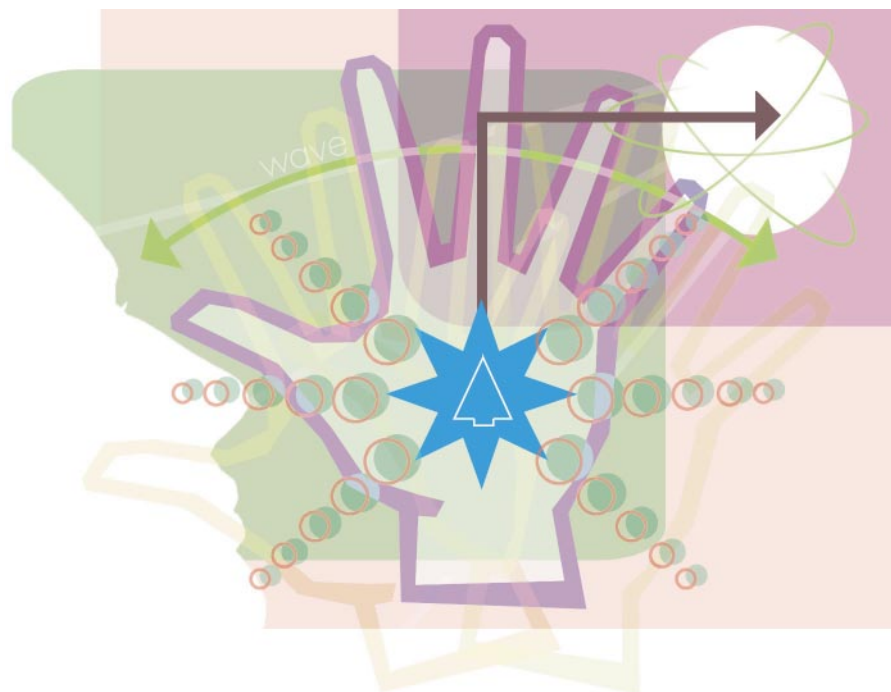
Humanity had become, by that autumn of 2350, two very distinct human species — if I can use that antiquated term. Oh, the Stock Humans remain a 'species' in the classical evolutionary sense: New People, of course, have forgone all that. Post-evolutionary, post-biological, budded or engineered, New People are gloriously free from all the old human restraints. What unites us all is our common source, the Divine Complexity that shaped primordial quark plasma into stars, planets, planaria, people. Grandfather taught me that.

I had always known that we would, one day, be separated. But we first spoke of it, tentatively and reluctantly, when Grandfather went with me to the Museum of Devices in Brussels, a day trip. I was young and easily impressed by the full-scale working model of a 'steam train' in the Machine Gallery — an amazingly baroque contrivance of ancient metalwork and gas-pressure technology. Staring at it, I thought (because Grandfather had taught me some of his 'religion'): Complexity made this. This is made of stardust, by stardust.

We walked from the Machine Gallery to the Gallery of the Planets, drawing more than a few stares from the Stock People (children, especially) around us. It was uncommon to see a New Person fully embodied and in public. The Great Goodbye had been going on for more than a century, but New People were already scarce on Earth, and a New Person walking with a Stock Person was an even more unusual sight — risqué, even shocking. We bore the attention gamely. Grandfather held his head high and ignored the muttered insults.

The Gallery of the Planets recorded humanity's expansion into the Solar System, and I hope the irony was obvious to everyone who sniffed at our presence there: Stock People could not have colonized any of these forbidding places (consider Ganymede in its primeval state!) without the partnership of the New. In a way, Grandfather said, this was the most appropriate place we could have come. It was a monument to the long collaboration that was rapidly reaching its end.

The stars, at last, are within our grasp.



The grasp, anyhow, of the New People. Was this, I asked Grandfather, why he and I had to be so different from one another?

"Some people," he said, "some families, just happen to prefer the old ways. Soon enough Earth will belong to the Stocks once again, though I'm not sure this is entirely a good thing." And he looked at me sadly. "We've learned a lot from each other. We could have learned more."

"I wish we could be together for centuries and centuries," I said.

I saw him for the last time (some years ago now) at the Shipworks, where the picturesque ruins of Detroit rise from the Michigan Waters, and the star-travelling Polises are assembled and wait like bright green baubles to lift, at last and forever, into the sky. Grandfather had arranged this final meeting — in the flesh, so to speak.

We had delayed it as long as possible. New People are patient: in a way, that's the point. Stock Humans have always dreamed of the stars, but the stars remain beyond their reach. A Stock Human lifetime is simply too short; one or two hundred years won't take you far enough. Relativistic constraints demand that travellers between the stars must be at home between the stars. Only New People have the continuity, the patience, the flexibility to endure and prosper in the Galaxy's immense voids.

I greeted Grandfather on the high embarkation platform where the wind was

brisk and cool. He lifted me up in his arms and admired me with his bright blue eyes. We talked about trivial things, for the simple pleasure of talking. Then he said, "This isn't easy, this saying goodbye. It makes me think of mortality — that old enemy."

"It's all right," I said.

"Perhaps you could still change your mind?"

I shook my head, no. A New Person can transform himself into a Stock Person and vice versa, but the social taboos are strong, the obstacles (family dissension, legal entanglements) almost insurmountable, as Grandfather knew too well. And in any case that wasn't my choice. I was content as I was. Or so I chose to believe.

"Well, then," he said, empty, for once, of words. He looked away. The Polis would be rising soon, beginning its eons-long navigation of our near stellar neighbours. Discovering, no doubt, great wonders.

"Goodbye, boy," he said.

I said, "Goodbye, Grandfather."

Then he rose to his full height on his many translucent legs, winked one dish-sized glacial blue eye, and walked with a slow machinelike dignity to the vessel that would carry him away. And I watched, desolate, alone on the platform with the wind in my hair, as his ship rose into the arc of the high clean noonday sky.

Robert Charles Wilson's novels include *Darwinia* and *Bios*, both published by Millennium.

A moving oxygen story

Harvey M. Flower

Titanium is one of the most versatile materials used in engineering. An innovative way of producing it from titanium dioxide will make the metal cheaper — if the process can be scaled up.

Imagine a metal that is as strong as many common engineering steels, yet more corrosion resistant than stainless steels and only little more than half as dense. The metal can be alloyed with a wide range of elements to enhance properties such as strength, toughness and creep resistance (the ability to resist deformation and fracture under stress at high temperatures). For good measure endow some of the alloys with a capacity for superplastic deformation, meaning that, under certain conditions, the metal can be stretched to many times its original length to form complex panel shapes. Finally, enable other alloys of the metal to form combinations that can remember their original shape after deformation and can return to it with a change of temperature or removal of stress.

This would be a remarkably attractive engineering material, and you do not in fact have to imagine it. It is titanium, and an imaginative application of the metal is shown in Fig. 1. But titanium is expensive, and on page 361 of this issue Chen and co-workers¹ describe an innovative way of producing the metal that could bring down the cost considerably. Another advantage is that their process avoids the use or production of the volatile and corrosive materials currently employed in titanium extraction.

At the moment, the total amount of titanium used annually worldwide is measured in tens of thousands of tonnes, whereas the consumption of steel is measured in many millions. The principal reason for this is cost, which can range from about \$7,500 per tonne for unalloyed titanium to \$40,000 or so per tonne for complex aerospace alloy products. The high price is not due to the rarity or difficulty of extracting titanium ores, which are plentiful and are the source of the titanium dioxide widely used as a white pigment in paint and other products. Rather, a large part (about one-third²) of the cost of titanium stems from the difficulty in reducing the oxide to the metal and ensuring that the residual oxygen content of the metal is both low and well controlled.

Conventional titanium extraction involves chlorinating titanium dioxide to produce a volatile, corrosive liquid, titanium tetrachloride. This is reduced to the metal by reaction with metallic magnesium (the Kroll process) or sodium (the Hunter process) at high temperature in sealed vessels. This batch process is slow, demanding in terms of



Figure 1 Titanium's sunlit glow: the Guggenheim Museum designed by Frank Gehry, in Bilbao, Spain. The outside surface of the museum (more than 30,000 square metres) is clad in corrosion-resistant titanium panels, which provide the building with a durable exterior and a striking visual appearance.

health and safety criteria, and requires prolonged high-temperature vacuum distillation of the product to remove excess alkali metal and its chloride from the titanium³.

The commercial history of titanium as an engineering material dates to about 1950. Ever since then, there have been attempts to develop alternative processes that lessen or eliminate the problems and cost of the chloride reduction route. Electrolysis, which is used commercially to extract aluminium, is potentially attractive. In the case of aluminium, the feedstock is aluminium oxide, which dissolves in the molten salt electrolyte (Na_2AlF_6 , cryolite). The product is molten and can be tapped off from the electrolysis cell, allowing continuous production.

In contrast, the electrolytic processes proposed for titanium have required conversion of the oxide to the volatile chloride, which is then dissolved in the molten electrolyte. Moreover, titanium becomes molten only at a temperature 1,000 K greater than that for aluminium, precluding the production of liquid metal. Removal of the solid metal deposited on the cathodes requires the cathodes to be periodically withdrawn from the bath and the titanium to be mechanically

stripped from them⁴. The process has not been the breakthrough in titanium production that had been hoped for.

Chen and co-workers¹ now propose an entirely novel electrolytic route, one that avoids conversion of the oxide to the chloride or the dissolution of the feedstock into the molten electrolyte. The authors provide convincing evidence that the reduction reaction proceeds through the ionization of oxygen from titanium dioxide at the cathode and its dissolution into the liquid calcium chloride electrolyte. Solid titanium is left behind. By controlling the electrode potential of the cathode, oxygen can even be removed from solution in the metal, allowing a high-purity product to be obtained. The resulting titanium is morphologically similar to the product of the Kroll process, being granular and porous. It can be removed as the reduced oxide cathodes themselves, or in particulate form if the walls of the electrolysis cell are used as cathodes and the oxide is introduced in pellet form.

Translating laboratory-scale experiments into industrial production inevitably poses problems of materials handling, and of designing and operating large electrolysis

YANN ARTHUS-BERTRAND/CO



100 YEARS AGO

As nest-building fishes are comparatively few, naturalists will read with interest an account given in the August issue of the *American Naturalist*, by Messrs. Young and Cole, of the manner in which the brook-lamprey (*Lampetra wilderi*) makes a structure of this nature. It is believed that the males precede the females at spawning time and commence nest-building before the arrival of the latter. The nest is made among pebbles, but it does not seem that the lampreys follow any definite plan in its construction. They affix themselves to such pebbles as require removing from the nest, and then endeavour to swim straight away with them. In the case of a heavy stone two lampreys may join forces. The number of fish in a nest may vary from one to thirty or forty; but there are generally between three and twenty-five.

From *Nature* 20 September 1900.

50 YEARS AGO

A sample of blood sent to us because of the presence of most unusual antibodies has proved, on investigation, to have even more extraordinary Rh antigens. The blood is unique in our experience, and in the literature, in that it has the antigen D, but lacks any detectable representative of the C and E allelomorph series of antigens. The donor of the blood is homozygous for this deficiency, owing without a doubt to her parents being second half cousins. Her mother is heterozygous for the condition; her father and two brothers are dead. The genotype of the donor may be written $-D-/-D-$, the dashes representing the absence of C, c, C', c', C'' and the absence of E, e, E'. The absence of these antigens was demonstrated in negative results of agglutination tests in saline and in albumin, of indirect anti-globulin tests, of trypsin tests and of absorption tests. ... In the serum of the donor we have so far been able to identify anti-e, anti-C and anti-c. The finding of both anti-C and anti-c in one serum is extraordinary but not altogether unexpected. ... Whatever the exact genetic mechanism, the fact that C and E are involved supports from an unexpected angle Fisher's tentative suggestion that the order of the genes on the chromosome would be found to be DCE. It also seems that the very controversial question of whether the genes are separable or not is settled in the most convincing way of all — by their separation.

From *Nature* 23 September 1950.

cells. However, the early results of trials being conducted at the Defence Evaluation and Research Agency, at Farnborough in Hampshire, offer hope that the process can be developed on a commercial scale and contribute to the cost reduction required to allow titanium to become a mainstream engineering material. Chen and colleagues also point out that the principle of oxygen ionization and removal by electrolysis can be applied to several other metal oxides. Here we have the intriguing possibility of producing pre-alloyed metals from mixed-oxide precursors.

Molecular biology

Small subunit, big science

James R. Williamson

Ribosomes are the cell's protein factories, and consist of two subunits, the large and the small. The structure of the small subunit of a bacterial ribosome has now been solved at atomic resolution, and is described in papers by Ramakrishnan and colleagues on pages 327 and 340 of this issue^{1,2}. These reports come shortly after the appearance of an independent, slightly lower-resolution structure of the small subunit³. Because this is the subunit where the genetic code is read, the new structural information shows in molecular detail how messenger RNAs (mRNAs) are faithfully translated into proteins.

The ribosome is a fundamental cellular component, found in both prokaryotes (loosely, eubacteria, which lack a nuclear membrane) and eukaryotes (which have such a membrane). It synthesizes all of the proteins in the cell, using mRNA as the template. The ribosome acts in concert with a variety of smaller factors which help to orchestrate the process, but the two main functions are entrusted to the ribosome itself: decoding the genetic code in the mRNA, and catalysing the formation of chemical bonds between amino acids, resulting in the polypeptide chains of proteins.

The large and small subunits of the eubacterial ribosome are known as 50S and 30S, respectively, from their rates of sedimentation in a centrifuge (the entire ribosome is 70S). A ribosome's catalytic site, the peptidyl transferase centre, is in the 50S subunit, and the structure of this subunit from the bacterium *Haloarcula marismortui* was described last month at a resolution of 2.4 Å (refs 4,5). Decoding of the mRNA occurs on the 30S subunit, and it is structural determination of 30S, to 3 Å from *Thermus thermophilus*, that is now described by Ramakrishnan and colleagues^{1,2}. Together, these papers

Harvey M. Flower is in the Department of Materials, Imperial College of Science, Technology and Medicine, London SW7 2BP, UK.
e-mail: h.flower@ic.ac.uk

1. Chen, G. Z., Fray, D. J. & Farthing, T. W. *Nature* **407**, 361–364 (2000).
2. Hartman, A. D., Gerdemann, S. J. & Hansen, J. S. *J. Mater.* **50**, 16–20 (1998).
3. Okura, Y. in *Titanium '95: Science and Technology* (eds Blenkinsop, P. A., Evans, W. J. & Flower, H. M.) Vol. 2, 1427–1437 (Inst. Materials, London, 1996).
4. Ginatta, M. V., Orsello, G., Perotti, P. & Berruti, R. in *Proc. Sixth World Conf. on Titanium* (eds Lacombe P., Tricot, R. & Beranger, G.) Vol. 2, 753–758 (Éditions de Physique, Zone Industrielle de Courtaboeuf, Les Ulis Cedex, 1998).

complement the earlier 7.8 Å structure of the entire 70S ribosome⁶.

The 30S subunit serves as the assembly guide for all of the factors needed in protein synthesis. The mRNA is a copy of the genomic 'repository' DNA sequence that encodes a gene product, and is bound to the 50S subunit in complex with transfer RNAs (tRNAs) via codon–anticodon interaction. The tRNAs are 'charged' with an amino acid, and deliver the appropriate amino acid specified by the codon. There are three tRNA binding sites on the ribosome, the A (acceptor), P (peptidyl) and E (exit) sites, each of which is occupied in succession by a particular tRNA

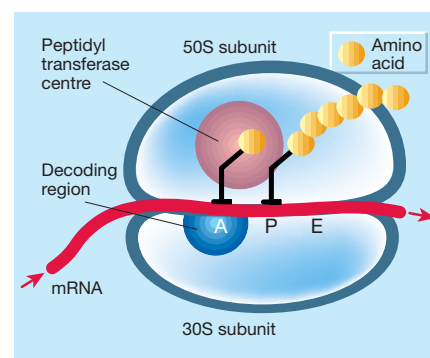


Figure 1 The two essential functions of the ribosome. The 30S (small) subunit contains the messenger RNA decoding site, and the 50S (large) subunit the peptidyl transferase centre. The mRNA threads through the subunit interface, and is decoded. The three transfer RNA binding sites, A (acceptor), P (peptidyl) and E (exit) then handle amino-acid selection, addition and completion of polypeptide synthesis. Here a charged tRNA is shown in the A site, and a nascent peptidyl-tRNA in the P site; the E site is vacant. Increasingly sophisticated structural characterization^{1–5} of the subunits and the ribosome as a whole will allow these events to be investigated in molecular detail.

during the protein synthesis cycle. The tRNAs bridge the large and small subunits, with the anticodon arm of the tRNA pointing towards the 30S subunit for decoding, and the acceptor arm of the tRNA pointing into the 50S subunit for peptidyl transferase (Fig. 1).

After protein synthesis begins, the growing peptide chain is covalently attached to the tRNA bound in the P site. The terminal amino group of the growing polypeptide is aimed into the A site, ready to link to the next amino acid to be delivered. The critical decoding event occurs as charged tRNAs bind to the A site and — in some unknown fashion — the cognate codon–anticodon interaction is sensed by the 30S subunit, triggering peptide-bond formation on the 50S subunit. The entire complex must then shift down the line to prepare for the next round of peptide-bond formation in the process known as translocation.

The new information on ribosome structure^{1–5} finally provides a molecular face to players that have long been known by their functions. The 30S subunit in particular is among the most intensively studied of cellular components, and a vast amount of genetic and biochemical information has accumulated about it over the past 40 years. Confronted with just the crystallographic coordinates, it would have been impossible to ascribe function to structural features. But this is not the case, and we can readily recognize the familiar functional sites. The locations of the A-, P- and E-site interactions with tRNA are clearly recognizable, and are in agreement with the biochemical and genetic data.

Furthermore, one of the papers² presents analysis of the 30S subunit in complex with several antibiotics. The significance of this is that antibiotics often work by inhibiting ribosome function, so blocking protein synthesis. Together with existing information about the effect of antibiotics on translation, the 30S antibiotic data provide insights into the mechanism of decoding and translocation.

The 30S subunit has a very different architecture to that of its 50S partner. In the 30S, clear domain boundaries are evident, and potentially flexible regions can be seen in cryo-electron microscopic reconstructions^{7,8}: large movements, on the scale of tens of angstroms, must occur on the 30S subunit during translocation of the mRNA from one codon to the next. In contrast, the 50S subunit is monolithic^{7,8}, its components being intricately folded and packed into a rigid structure. These unlikely colleagues — the squat and stolid 50S subunit and the lithe and flamboyant 30S subunit — work together to carry out one of the most intricate functions of all cells.

It is perhaps not widely appreciated that about two-thirds of the ribosome's mass is composed of RNA. With atomic knowledge

of the 30S and 50S structures, and how they assemble into the overall 70S ribosome, one point stands out — the most essential functions of the ribosome are carried out by RNA⁹. Historically, biologists have concentrated on genes and proteins. The dogma that nucleic acids are the repository of information and that proteins catalyse chemical reactions was overturned only 15 years ago by the demonstration of catalytic RNAs^{10,11}. Despite increasing support for the RNA-world hypothesis, which sees RNA-based life-forms as the progenitors of modern cells, most biologists did not seriously consider the possibility that RNA could be playing more than a bit part in reactions such as protein synthesis. Well, wake up and smell the coffee. In the ribosome, it has turned out that most of the intersubunit interface is RNA⁶, the peptidyl transferase centre is RNA^{4,5,12}, and the decoding site and most of the A and P sites are RNA^{1–3}. It appears that the modern ribosome is composed of a somewhat geriatric, but functionally vital, RNA scaffold

that is propped up and doted upon by its protein grandchildren. The ribosome is one colossal RNA enzyme. ■

James R. Williamson is in the Department of Molecular Biology and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA.

e-mail: jrwill@scripps.edu

1. Wimberly, B. T. *et al.* *Nature* **407**, 327–339 (2000).
2. Carter, A. P. *et al.* *Nature* **407**, 340–348 (2000).
3. Schlunzen, F. *et al.* *Cell* **102**, 615–623 (2000).
4. Ban, N., Nissen, P., Hansen, J., Moore, P. B. & Steitz, T. A. *Science* **289**, 905–920 (2000).
5. Nissen, P., Hansen, J., Ban, N., Moore, P. B. & Steitz, T. A. *Science* **289**, 920–930 (2000).
6. Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. *Science* **285**, 2095–2104 (1999).
7. Gabashvili, I. S. *et al.* *EMBO J.* **18**, 6501–6507 (1999).
8. Gabashvili, I. S., Agrawal, R. K., Grassucci, R. & Frank, J. J. *Mol. Biol.* **286**, 1285–1291 (1999).
9. Cech, T. R. *Science* **289**, 878–879 (2000).
10. Zaug, A. J. & Cech, T. R. *Cell* **19**, 331–338 (1980).
11. Cech, T. R., Zaug, A. J. & Grabowski, P. J. *Cell* **27**, 487–496 (1981).
12. Muth, G. W., Ortoleva-Donnelly, L. & Strobel, S. *Science* **289**, 947–950 (2000).

Astronomy

Galactic rotation in real time

John Kormendy

On page 349 of this issue, Andrea Ghez and colleagues¹ present the first measurements of the acceleration of stars orbiting the nucleus of our Galaxy. These stars are part of a dense cluster associated with the compact radio source known as Sagittarius A*, which is thought to be at the actual centre of the Milky Way. Measuring the accelerations of stars in galactic orbits (and not just their velocities) is a technical tour de force. These accelerations are enormous on a Galactic scale; at 3–6 mm s⁻² they are comparable to those experienced by the Earth as it orbits the Sun. With these measurements, the authors help to identify Sagittarius A* as the supermassive black hole at the centre of the Milky Way. They also test — and provide support for — fundamental assumptions about the dynamics of the Galactic centre and our understanding of violent activity in galactic nuclei.

Ghez *et al.*¹ took a series of infrared images of the nucleus of the Milky Way with the 10-metre Keck telescope on Mauna Kea, Hawaii. The Galactic centre is 8 kiloparsecs away (1 parsec = 3.26 light years) and Ghez *et al.* study an area 0.2 parsecs across. By taking short exposures they freeze the jitter in the image caused by turbulence in the Earth's atmosphere and produce images with 0.05 arcsec (0.002 parsecs) resolution. From such an image, one can accurately determine stellar positions. The stars nearest the centre move significantly in just a few years, so Ghez

et al. can measure the two components of their velocities that are in the plane of the sky. If the central star cluster is in equilibrium, then its dynamics tell us the total mass enclosed inside it.

Past velocity measurements^{2–5} have already established that a dark mass equal to 2.6×10^6 solar masses ($M_{\odot}$) lies in the vicinity of Sagittarius A*. For the first time, Ghez *et al.* provide acceleration measurements for three stars within 0.005–0.013 parsecs of Sagittarius A* (Fig. 1, overleaf). Acceleration vectors provide important new information about the dark mass by pointing to the location of the gravity source. The acceleration vectors are still fairly uncertain — observations span only 4 years — but they intersect within 0.002 ± 0.0016 parsecs of Sagittarius A*. This strengthens the association of Sagittarius A* with the dynamically detected dark mass — an object with enough mass to bend the orbits of stars moving at speeds of up to 1,350 km s⁻¹.

Sagittarius A* is widely believed by astronomers to be the Galaxy's central black hole. It is appropriately tiny — similar in radius to the orbit of Mars around the Sun⁶. Nevertheless, for a black hole of a few million solar masses, it is surprisingly dark. The strong gravity of a black hole prevents radiation escaping from inside a certain radius (the Schwarzschild radius). But gas that is still falling toward the point of no return is accelerated to almost the speed of light in an

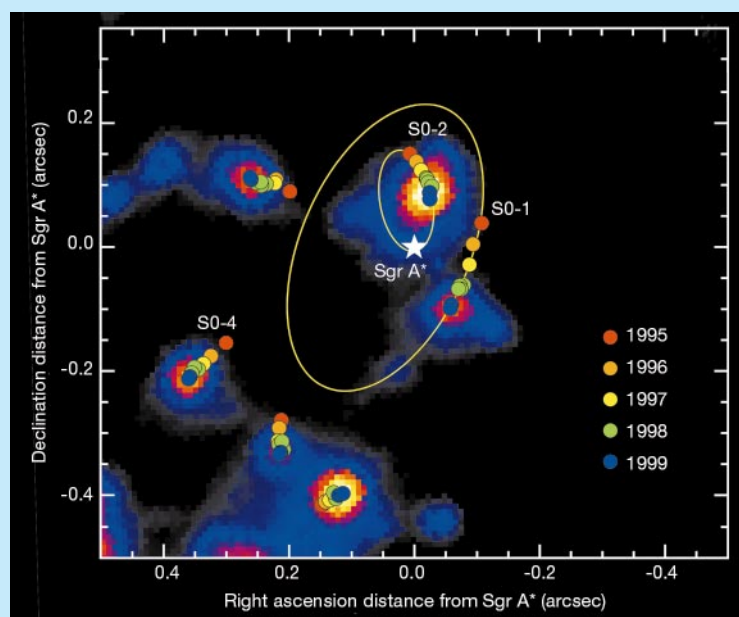


Figure 1 Infrared image of the Sagittarius (Sgr) A* star cluster taken in May 1999. Filled circles show the positions of six stars at different epochs. Ghez *et al.*¹ have measured stellar accelerations for three stars close to the Galactic centre (S0-1, S0-2 and S0-4). The orbits of S0-1 and S0-2 (yellow ellipses) were calculated by assuming that Sagittarius A* is a supermassive black hole with a mass of 2.6×10^6 solar masses. The orbital periods are 63 and 17 years, respectively; they are near the short end of the range of orbits allowed by the acceleration measurements. (Image kindly provided by A. Ghez.)

'accretion disk' that radiates ferociously. This process is believed to power active galactic nuclei, such as quasars. In our Galaxy, fuel is plentiful: hot, young stars near the Galactic centre blow a wind of gas past the black hole. Why, then, is it not brighter?

The radiation that we see at the centre of the Milky Way is believed to be the faint squeaking of a black hole that is hiding a substantial diet by advecting and swallowing the energy in its accretion disk before it has time to be radiated away⁷. So, although the suspected black hole at the centre of our Galaxy has become a cornerstone of the theory^{8,9} that violent activity in galactic nuclei is powered by accretion onto black holes, the evidence that the central dark mass really is a black hole is very indirect. Accordingly, substantial effort has gone into testing alternative interpretations of the observations.

Black holes are extraordinarily slippery. The stars that Ghez *et al.* use in their analysis are 50,000 Schwarzschild radii from the centre. This is closer to the centre than in any other candidate black hole, but it is far outside the region of strong gravity. That is, the largest stellar velocity, $1,350 \text{ km s}^{-1}$, is much smaller than the speed of light. So the arguments for a black hole are still indirect. The most plausible alternative in astrophysical terms is a clump of dark stars — stars that are too low in mass to ignite nuclear reactions or stars that have died and faded from sight. Past measurements have determined a small enough radius for the dark mass to exclude

these alternatives^{3,4,10}. Ghez *et al.* strengthen these arguments by finding a new minimum density of the dark mass, $8 \times 10^{12} M_{\odot} \text{ parsec}^{-3}$, that is almost an order of magnitude larger than previously found. Squeezing dark stars to such a high density has consequences that are excluded by the observations.

Other more exotic possibilities remain. One example is a hypothetical ball of neutrinos^{11,12}. But a neutrino ball of $2.6 \times 10^6 M_{\odot}$ is

expected to have a radius of about 0.02 parsecs. Two of the three stars studied by Ghez and colleagues are well inside this radius, only 0.005 parsecs from Sagittarius A*. The future paths of these stars depend sensitively on whether there is a black hole or a neutrino ball at the centre of the Galaxy. So we can test some of the more exotic black-hole alternatives by following the stars as they move through substantial orbit arcs. This should be possible within a few more years.

Those of us who study galactic dynamics generally observe a snapshot of objects that change only on timescales of millions or billions of years. The orbital periods of the stars measured by Ghez *et al.* can be as short as a few decades. There is something quite grand in the realization that we can expect, with good health and a little luck, to see the Galactic centre rotate at least once in our lifetimes.

John Kormendy is in the Department of Astronomy, University of Texas at Austin, RLM 15.308, Austin, Texas 78712, USA.

e-mail: kormendy@astro.as.utexas.edu

- Ghez, A. M., Morris, M., Becklin, E. E., Tanner, A. & Kremenek, T. *Nature* **407**, 349–351 (2000).
- Eckart, A. & Genzel, R. *Mon. Not. Astron. Soc.* **284**, 576–598 (1997).
- Genzel, R., Eckart, A., Ott, T. & Eisenhauer, F. *Mon. Not. R. Astron. Soc.* **291**, 219–234 (1997).
- Genzel, R., Pichon, C., Eckart, A., Gerhard, O. E. & Ott, T. <http://xxx.lanl.gov/abs/astro-ph/0001428>
- Ghez, A. M., Klein, B. L., Morris, M. & Becklin, E. E. *Astrophys. J.* **509**, 678–686 (1998).
- Lo, K. Y., Shen, Z.-Q., Zhao, J.-H. & Ho, P. T. P. *Astrophys. J.* **508**, L61–L64 (1998).
- Narayan, R., Mahadevan, R. & Quataert, E. in *The Theory of Black Hole Accretion Discs* (eds Abramowicz, M. A., Björnsson, G. & Pringle, J. E.) 148–182 (Cambridge Univ. Press, 1998).
- Rees, M. J. *Annu. Rev. Astr. Astrophys.* **22**, 471–506 (1984).
- Blandford, R. D. in *Active Galactic Nuclei: Saas-Fee Course 20* (eds Courvoisier, T. J.-L. & Mayor, M.) 161–275 (Springer, Berlin, 1990).
- Maoz, E. *Astrophys. J.* **494**, L181–L184 (1998).
- Tsiklauri, D. & Viollier, R. D. *Astrophys. J.* **500**, 591–595 (1998).
- Munyanza, F., Tsiklauri, D. & Viollier, R. D. *Astrophys. J.* **526**, 744–751 (1999).

Cancer

A radical approach to treatment

John L. Cleveland and Michael B. Kastan

Reactive oxygen species are potentially dangerous by-products of cellular metabolism that have direct effects on cell development, growth and survival, on ageing, and on the development of cancer¹. They are generated by all aerobic organisms, but their production is a double-edged sword. On the one hand, they seem to be needed for signal-transduction pathways that regulate cell growth² and reduction–oxidation (redox) status¹. But on the other, excessive amounts of these metabolites can start lethal chain reactions, which oxidize and disable structures that are required for cellular integrity and survival¹. Many

tumour cells seem to have increased rates of metabolism compared with normal cells, which would typically lead to increased numbers of reactive oxygen species. So one way of treating cancer might be to design drugs to target the enzymes that regulate the levels of reactive oxygen species. On page 390 of this issue³, Huang and colleagues provide support for this idea: they find that a compound that kills human leukaemia cells, but spares normal cells, may achieve this by inhibiting such an enzyme.

Reactive oxygen species are generated during the production of ATP by aerobic metabolism in mitochondria. The leakage of

electrons from mitochondria during the electron-transport steps of ATP production generates the reactive oxygen species superoxide (O_2^-) and hydroxyl ($OH^\bullet$) radicals. These species can lead to the production of hydrogen peroxide (H_2O_2), from which further hydroxyl radicals are generated in a reaction that either depends on, or is catalysed by, Fe^{2+} ions¹ (Fig. 1). Cells have evolved a series of antioxidant systems to handle these dangerous natural by-products. These defence systems include intracellular superoxide dismutases (SODs), which convert O_2^- into H_2O_2 ; enzymes that inactivate H_2O_2 or hydroxyl radicals; and enzymes that trap free radicals or transition metals (such as Fe^{2+}) that are a reservoir for electrons.

A higher level of aerobic metabolism — and hence of reactive oxygen species — in tumour cells compared with normal cells would make enzymes in these pathways good candidate drug targets. However, previous attempts to inhibit SODs have not proved clinically useful. For example, ions that successfully compete with O_2^- to bind to SODs were not specific enough to cancer cells and produced unacceptable levels of damage to normal cells. But Huang *et al.*³ now show that a group of oestrogen derivatives, of all things, seem to bind to and inhibit SODs, and that — by damaging mitochondria — these compounds induce the suicide of a variety of human leukaemia cells *in vitro*.

Huang *et al.*'s insights are their reward for their persistent investigation of the observation that the oestrogen derivative in question — 2-methoxyoestradiol — induces the death of cancer cells. This oestrogen derivative is a normal product of the metabolism of oestradiol (a type of oestrogen), and has a rather rich history in the literature. Although it is an oestrogen derivative, 2-methoxyoestradiol cannot bind to the oestrogen receptor⁴. But it does induce the suicide of cancer cells *in vitro* and *in vivo*^{5–7}. This result has been variously attributed to the effects of 2-methoxyoestradiol on the formation of blood vessels to feed the tumour, the polymerization of tubulin (a component of the cell's cytoskeleton), and/or the expression of cell-suicide and cell-division regulators^{5–7}. The findings of Huang *et al.*³ indicate that these different effects may in fact result from the inhibition of SODs.

There are two intracellular SODs: the mitochondrial-associated manganese-dependent SOD (MnSOD), and the cytoplasmic/nuclear copper/zinc-dependent SOD (CuZnSOD). The authors use a variety of techniques to show that the expression of CuZnSOD is induced by 2-methoxyoestradiol in leukaemia cells. They reason that this could reflect one of two scenarios. First, 2-methoxyoestradiol might cause an increase in the amount of dangerous O_2^- in the tumour cells; more SODs would therefore be needed to mop up the radicals. If this were

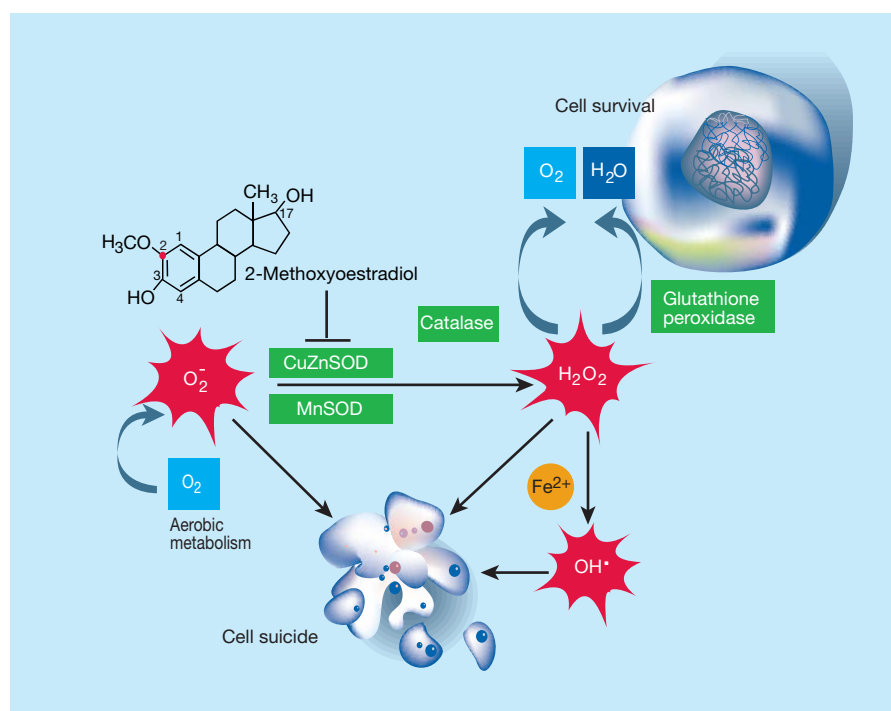


Figure 1 Inhibition of superoxide dismutases triggers the suicide of leukaemia cells. Superoxide (O_2^-), a toxic free radical, is a cellular by-product of aerobic metabolism. Superoxide is mopped up by two superoxide dismutases (SODs) — copper/zinc-dependent SOD (CuZnSOD) and manganese-dependent SOD (MnSOD) — and then by other enzymes such as catalase and glutathione peroxidase. This allows the survival of the cell. Cell suicide can also occur if hydrogen peroxide (H_2O_2) levels rise too much or if H_2O_2 is instead converted to hydroxyl radicals ($OH^\bullet$) in an Fe^{2+} -dependent reaction¹. Cancer cells have high levels of metabolism that generate reactive oxygen species, and may sometimes have low levels of SOD activity. Huang *et al.*³ show that treating leukaemia cells with the oestradiol derivative 2-methoxyoestradiol leads to increases in the expression of CuZnSOD, and that this reflects the cell's response to inhibition of SOD activity by 2-methoxyoestradiol. The result is a net increase in superoxide levels and the suicide of leukaemia cells. A derivative of this protein with a hydroxyl substitution at the 2-carbon (red) can also kill leukaemia cells.

true, then more H_2O_2 should be produced in the leukaemia cells in response to 2-methoxyoestradiol (Fig. 1). But instead, Huang *et al.* find that 2-methoxyoestradiol causes a dose-dependent decrease in H_2O_2 levels. The second possibility, which fits with this observation, is that 2-methoxyoestradiol suppresses SOD activity. This might sound counterintuitive, but if this were to happen, the cells would sense this change and respond by increasing the level of CuZnSOD.

Indeed, the authors show that 2-methoxyoestradiol effectively inhibits the activity of both CuZnSOD and MnSOD *in vitro*, and binds to SODs in intact cells. Huang *et al.* also find that if they manipulate SOD levels, they can control the sensitivity of ovarian cancer cells to 2-methoxyoestradiol. And O_2^- scavengers or antioxidants protect leukaemia cells from the lethal effects of this compound.

These are provocative findings, but they raise several questions. First, is SOD the only relevant target of 2-methoxyoestradiol that is required for the suicide of leukaemia cells? Second, we need to know more about how this compound inhibits SOD. For example,

where does it bind, and does it prevent the binding of O_2^- to SODs? Structural studies should help in answering these questions. Third, these findings imply that this and similar oestrogen derivatives (Fig. 1) — or other compounds that inhibit SODs — could be used broadly in cancer treatment. But, as Huang *et al.* point out, there could be cell-specific circumstances in which inhibiting SODs might compromise tumour treatment. Finally, will 2-methoxyoestradiol or similar oestrogen analogues truly be useful for treatment purposes? For example, these compounds might have toxic side effects: 2-methoxyoestradiol compromises the growth of new blood cells^{5,7} (which is needed for normal as well as tumour cells) and can damage DNA *in vitro*⁸.

Whatever the answers, the results³ raise hope that inhibiting SODs at the same time as treating cells with agents that increase the levels of reactive oxygen species is a promising way of treating at least some cancers. At the very least, even if SODs turn out not to be ideal drug targets, we can now be confident that it is reasonable to try to exploit the differences in the ways that normal and tumour cells control their redox status.

John L. Cleveland and Michael B. Kastan are in the Departments of Biochemistry and Hematology/Oncology, St Jude Children's Research Hospital, 332 N. Lauderdale, Memphis, Tennessee 38018, USA. e-mails: john.cleveland@stjude.org michael.kastan@stjude.org

1. Davies, K. J. *Biochem. Soc. Symp.* **61**, 1–31 (1995).
2. Sundaresan, M., Yu, Z.-X., Ferrans, V. J., Irani, K. & Finkel, T. *Science* **270**, 296–299 (1995).

3. Huang, P., Feng, L., Oldham, E. A., Keating, M. J. & Plunkett, W. *Nature* **407**, 390–395 (2000).
4. Cushman, M., He, H. M., Katzenellenbogen, J. A., Lin, C. M. & Hamel, E. *J. Med. Chem.* **38**, 2041–2049 (1995).
5. Fotsis, T. *et al. Nature* **368**, 237–239 (1994).
6. Mukhopadhyay, T. & Roth, J. A. *Oncogene* **14**, 379–384 (1997).
7. Klauber, N., Parangi, S., Flynn, E., Hamel, E. & D'Amato, R. J. *Cancer Res.* **57**, 81–86 (1997).
8. Tsutsui, T. *et al. Carcinogenesis* **21**, 735–740 (2000).

Ocean biogeochemistry

Calcification and CO₂

Jean-Pierre Gattuso and Robert W. Buddemeier

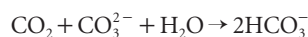
Most concerns about rising concentrations of atmospheric carbon dioxide centre on how climate may change. But there may also be direct biological effects. In terrestrial ecosystems, extra atmospheric CO₂ may have a fertilizing effect, resulting in increased photosynthesis. Except for a few organisms, such as seagrasses, such an increase seems unlikely in marine systems because most algae use bicarbonate ions (HCO₃[−]) rather than CO₂ as a photosynthetic substrate. There have, however, been indications of a negative biological response in the marine environment. Several studies^{1–4} have shown that calcification rates of reef-building corals and coralline algae are depressed by increased levels of CO₂. On page 364 of this issue, Riebesell *et al.*⁵ considerably extend the significance of these earlier results by describing the influence of CO₂ on the calcification of coccolithophorids — an important group of algae widely distributed in the surface layer of coastal waters and the open ocean.

The authors provide experimental evidence that the calcification rate of two species of coccolithophorids maintained in culture is lower at CO₂ levels predicted for 2100 than at preindustrial levels. Natural communities that the authors sampled in the North Pacific showed a similar response. Coccolithophorids are distributed throughout the world's oceans (Fig. 1), and may be the world's most important producer of calcium carbonate (CaCO₃). They form low-magnesium calcite shells and their distribution extends into subpolar waters.

The reef-building corals and coralline algae known to be sensitive to high CO₂ levels^{1–4} are bottom-dwelling. They occur mostly in tropical and subtropical regions, although some of the algae extend to high latitudes, and they typically form skeletons of aragonite or high-magnesium calcite. Together, the coccolithophorids, reef-building corals and coralline algae are responsible for well over half of the world's CaCO₃ production⁶. Given their widespread combined distribution, it seems likely that the carbon-

ate biogeochemistry, and possibly the ecology, of the entire world ocean will be altered by rising CO₂ levels (Fig. 1).

Atmospheric CO₂ equilibrates rapidly with the surface layer of the ocean, where most additional CO₂ combines with carbonate ions:



This leads to a decrease in the concentration of CO₃^{2−}, one of the building blocks of calcium carbonate, and in the saturation state of calcium carbonate, Ω ($\Omega = [\text{Ca}^{2+}] \times [\text{CO}_3^{2-}] / K_{sp}$, where K_{sp} is the equilibrium constant of CaCO₃). Rather than CO₂, Ω seems to be the controlling factor, because increasing Ω by adding bicarbonate promotes calcification even though it induces higher concentrations of dissolved CO₂ (ref. 2).

Decreased calcification in response to increased concentrations of CO₂ has now been reported across phylogenetically distant groups that precipitate different types of CaCO₃ crystals (low-magnesium calcite,

high-magnesium calcite and aragonite), either through internal or external calcification. But the unifying theme is photosynthesis: to date all of the calcifying organisms shown to have a strong saturation-state response are plants or animals that depend on photosynthetic symbionts.

Coccolithophorids are widely used as a model organism to investigate photosynthesis and calcification. Both processes, and their interaction, are relatively well understood in these algae. We know little, however, of how they might respond to environmental change. Riebesell and colleagues⁵ have not only begun to fill that gap, but their results also draw attention to the likelihood that the critical factor in responses to higher CO₂ is not as much the mineralogy as the link to photosynthesis.

Rising CO₂ levels have two antagonistic effects on the CO₂ fluxes mediated by calcification. The amount of CO₂ generated by calcification will fall as a result of the decreased rate of CaCO₃ precipitation^{1,5}. However, increased CO₂ concentration also shifts the seawater carbonate equilibria with the result that more CO₂ is released per mole of CaCO₃ precipitated⁷. These two effects could be balanced in coral reefs¹ but the response of pelagic calcification could lead to increased CO₂ storage capacity in the upper ocean⁵. CaCO₃ compensation is the CaCO₃ dissolution that maintains a balance between inputs from land and burial in deep-ocean sediments; this process is a potential regulator of atmospheric CO₂ during glacial–interglacial cycles⁸, and could also be affected. Dissolution is favoured by decreasing Ω and increasing coccolithophorid primary production relative to calcification⁵, the so-called rain ratio.

What are the main issues to be resolved? First, the role of photosynthetically coupled calcification in the global carbonate cycle,

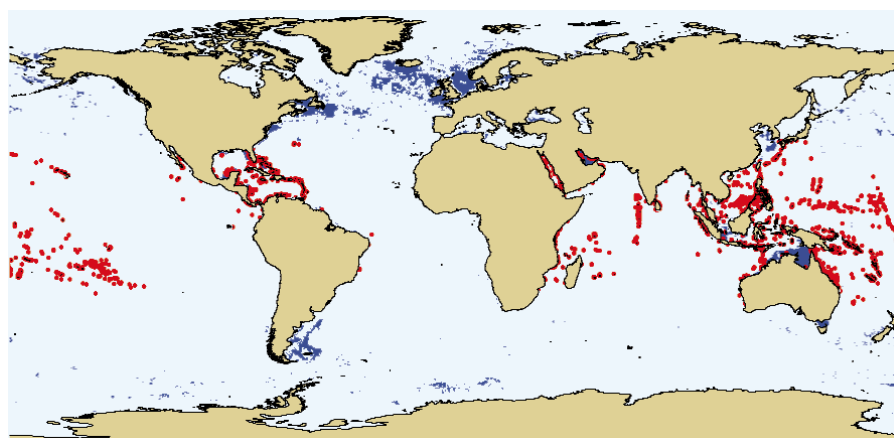


Figure 1 Distribution of coccolithophorids and coral reefs, the major photosynthetic and calcifying systems in the marine environment. Coccolithophorid blooms (blue dots) are transient and annually cover $14 \times 10^5 \text{ km}^2$, mostly in the subpolar regions¹¹. Coral reefs (red dots) are permanent, long-lived ecosystems that cover $6 \times 10^5 \text{ km}^2$ in the tropics¹². Taken together, several studies^{1–5}, including that now reported by Riebesell *et al.*⁵, show that the carbonate biogeochemistry of the world's oceans is subject to alteration by rising levels of CO₂. (Distribution of coccolithophorid blooms is from ref. 11.)

and its sensitivity to ocean chemistry, requires further exploration. Second, marine biological responses to high levels of CO₂ have been investigated in the short-term; but the CO₂ level expected in 2100 will be the highest for at least the past 24 million years⁹, and physiological acclimatization and genetic adaptation might, at least in part, counteract adverse effects. Third, interactions of CO₂ with other environmental changes such as increased temperature and nutrient concentrations need to be investigated.

Prediction of the response of ocean chemistry to rising CO₂ is reasonably straightforward¹⁰. But long-term predictions of the biological responses to climate and environmental changes, and of their feedback on climate in the future, are hampered by our limited understanding of the interactions concerned and by the modest database on which models and predictions are based. To tackle these deficiencies, several programmes are either underway or in the final stages of planning. Two of them are the US Carbon Cycle Science Plan and the international Surface Ocean Lower Atmosphere

Study (SOLAS), both of which can potentially resolve at least some of the outstanding questions.

Jean-Pierre Gattuso is at the Observatoire Océanologique, Laboratoire d'Océanographie Biologique et Écologie du Plancton Marin, CNRS-UPMC, BP 28, 06234 Villefranche-sur-Mer Cedex, France.

e-mail: gattuso@obs-vlfr.fr

Robert W. Buddemeier is at the Kansas Geological Survey, University of Kansas, 1930 Constant Avenue, Lawrence, Kansas 66047, USA.

e-mail: buddrw@kgs.ukans.edu

1. Gattuso, J.-P., Allemand, D. & Frankignoulle, M. *Am. Zool.* **39**, 160–183 (1999).
2. Marubini, F. & Thake, B. *Limnol. Oceanogr.* **44**, 716–720 (1999).
3. Langdon, C. *et al. Glob. Biogeochem. Cycles* **14**, 639–654 (2000).
4. Leclercq, N., Gattuso, J.-P. & Jaubert, J. *Glob. Change Biol.* **6**, 329–334 (2000).
5. Riebesell, U. *et al. Nature* **407**, 364–367 (2000).
6. Milliman, J. D. & Droxler, A. W. *Geol. Rund.* **85**, 496–504 (1996).
7. Frankignoulle, M., Canon, C. & Gattuso, J.-P. *Limnol. Oceanogr.* **39**, 458–462 (1994).
8. Archer, D., Winguth, A., Lea, D. & Mahowald, N. *Rev. Geophys.* **38**, 159–189 (2000).
9. Pearson, P. N. & Palmer, M. R. *Nature* **406**, 695–699 (2000).
10. Kleypas, J. A. *et al. Science* **284**, 118–120 (1999).
11. Brown, C. W. & Yoder, J. A. *J. Geophys. Res.* **99**, 7467–7482 (1994).
12. Smith, S. V. *Nature* **273**, 225–226 (1978).

presented on the cell surface or recognized by the CTLs. In theory, this would enable HIV to escape from CTLs, contributing to the persistence of HIV in the body and to the progression of the disease. HIV can escape from these T cells during an infection^{4–6}, but this has not been consistently seen. Part of the difficulty in answering the question of how important viral variation and escape from CTLs are in AIDS progression lies in the fact that the initial DNA and protein sequences of the infecting strain of virus are usually not known⁷.

Another problem with the earlier studies was that, in most of them, researchers looked at CTL responses to single HIV peptides that were identified during progressive, uncontrolled infection rather than the early, acute stages. These peptides include those from the proteins Gag, reverse transcriptase, Env and Nef. But, because early immune events have a strong influence on immune control over HIV and the speed of AIDS progression^{8,9}, the CTL responses during the acute phase of infection are perhaps of greater interest.

Allen *et al.*² now fill in several of the gaps in our knowledge. The authors injected an SIV strain into rhesus macaques, after determining the precise DNA and protein sequences of the virus. This meant they could track the changes that occurred in the DNA and proteins of the virus during the course of an infection. They also looked at the CTL responses to so-called viral regulatory proteins (which include Tat) — among the earliest proteins produced in infected cells — as well as to Gag and other SIV proteins.

The authors found that the number of SIV-specific CTLs peaked after three to four weeks of infection. Most of these T lymphocytes recognized peptides from either Gag or Tat. Levels of virus in the blood declined as these specific CTLs appeared, but SIV per-

AIDS

Escape from the immune system

Bruce D. Walker and Philip J. R. Goulder

One of the key unanswered questions about human infection with HIV is why the immune system ultimately fails to control the virus. Cytotoxic T lymphocytes (CTLs) can kill HIV-infected cells and limit the production of new viruses¹, but the level of control achieved in HIV infection is far less than in other viral infections, and is not enough to prevent the disease from progressing further.

On page 386 of this issue², Allen and colleagues use a monkey model of AIDS to show that, in the early stages of infection, many of the CTLs generated locate infected cells by recognizing a viral protein called Tat. These Tat-specific killer cells are very efficient at controlling the strain of virus that first infects an individual. But mutations within Tat arise while this immune response is developing, and the mutant viruses escape detection, establishing a long-lasting infection. The results raise the possibility that vaccines that induce the production of a mature, Tat-specific T-cell response might be an effective way of giving the human immune system a head start in the fight against HIV infection.

Many groups³ have looked at the response of CTLs to HIV and its counterpart in monkeys, simian immunodeficiency virus (SIV). But these studies did not resolve why

the killer cells are incapable of adequately controlling viral number. CTLs recognize an HIV-infected cell by detecting viral peptides — typically eight to ten amino acids long — that are processed by the infected cell and displayed on its surface. But even a single amino-acid change in the viral peptide can be enough to prevent the peptide from being

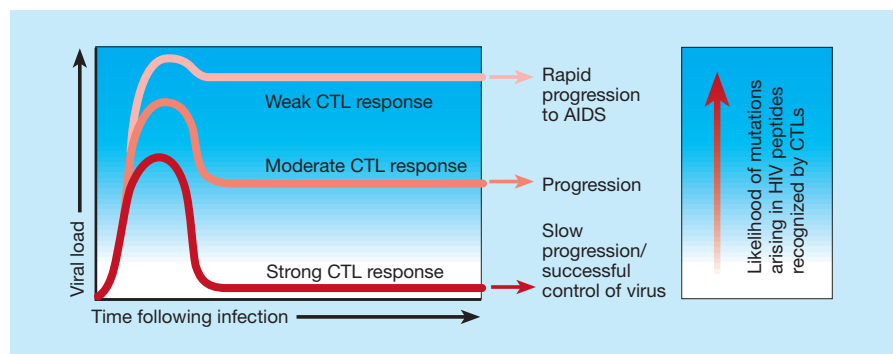


Figure 1 The immune system's response to AIDS-causing viruses. Cytotoxic T lymphocytes (CTLs) seek out virus-infected cells and destroy them, reducing virus spread. They do this by detecting viral peptides, which are processed by an infected cell and displayed on its surface. But in general CTLs are unable to control virus numbers well enough to prevent the progression to AIDS. As shown by Allen *et al.*², this failure results in part from the fact that AIDS viruses can mutate to evade recognition by effective CTLs, such as the early CTL response directed towards the Tat protein. Increasing the efficiency of the early CTL response by vaccination before exposure to HIV may be an effective way of improving the initial control over the virus, reducing the likelihood that mutations will arise in important CTL peptides such as those in Tat.

sisted nonetheless. The authors sequenced the virus DNA and proteins after eight weeks, and found that none of the initial SIV strain remained. The viruses had evolved from that first strain, and were able to escape being recognized by the CTLs. Changes occurred in the Tat peptide in all ten of the animals studied in detail. The rest of the virus appeared to be largely untouched by immune pressure. The authors' finding that the steady-state viral load was lower in animals with early, Tat-specific CTL responses attests to the important antiviral effect of these cells.

One of the most intriguing questions raised by these results is why the virus escaped by mutation in the Tat peptide, rather than the Gag peptide, given that the timing and magnitude of the CTL responses to these two peptides seemed to be similar. Perhaps Tat-specific CTLs are more effective at controlling the virus, and so impose a greater selection pressure on the virus. Alternatively, sequence changes may be better tolerated in certain regions of the viral proteins but constrained in others. For example, the Tat peptide is derived from a region of the Tat protein with no defined function, so it might not be damaging to the protein if this peptide were allowed to mutate.

These results have implications in the quest for an HIV vaccine. First, we can no longer look at CTL responses to a single viral peptide as a substitute for overall CTL function — instead, we need to consider the breadth of the immune response. Second, the results suggest that all CTL responses are not created equal, so we need to determine whether CTL specific for Tat and the other regulatory proteins are the most effective at controlling HIV infection. Finally, particu-

larly in studies of acute infection, we need to look at the CTL responses to the infecting strain of virus, as opposed to a consensus, published sequence. The greater effort required to sequence the infecting strain will no doubt be rewarded by a much better understanding of where immune pressure is being applied.

So it seems that Tat-specific CTLs are a key part of the immune system's early response to SIV, raising the possibility that they might be important in the development of an HIV vaccine. Other results indicate that using Tat may be a promising way to stimulate the immune system¹⁰. A pre-existing Tat-specific immune response at the time of virus infection would be expected to inhibit virus more effectively than normal, and to prevent the high level of HIV replication that promotes mutations and immune escape (Fig. 1). This idea can now be tested by experimentally inducing the production of Tat-specific CTLs in macaques, and then introducing SIV. The value of animal models of disease is perhaps nowhere more clearly seen than in these studies².

Bruce D. Walker and Philip J. R. Goulder are at the Partners AIDS Research Center, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts 02114, USA.
e-mail: bwalker@helix.mgh.harvard.edu

1. Yang, O. O. *et al.* *J. Virol.* **71**, 3120–3128 (1997).
2. Allen, T. M. *et al.* *Nature* **407**, 386–390 (2000).
3. Goulder, P. J. R. *et al.* *AIDS* **13** (suppl. A), S121–S136 (1999).
4. Goulder, P. J. R. *et al.* *Nature Med.* **3**, 212–217 (1997).
5. Borrow, P. *et al.* *Nature Med.* **3**, 205–211 (1997).
6. Price, D. A. *et al.* *Proc. Natl Acad. Sci. USA* **94**, 1890–1895 (1997).
7. Brander, C. *et al.* *J. Clin. Invest.* **101**, 2559–2566 (1998).
8. Pantaleo, G. *et al.* *Proc. Natl Acad. Sci. USA* **94**, 254–258 (1997).
9. Lyles, R. H. *et al.* *J. Infect. Dis.* **181**, 872–880 (2000).
10. Cafaro, A. *et al.* *Nature Med.* **5**, 643–650 (1999).

epidermis (the outermost cell layer) and in shaping appendages. One long-standing mystery about Wingless activity in the epidermis is the asymmetric distribution of Wingless protein in each epidermal segment. Previous studies indicated that a barrier prevents the spread of Wingless to posterior parts of each segment¹. But now it seems that the asymmetry may result instead (or also) from cells in different parts of the segment having differing abilities to degrade Wingless protein (J.-P. Vincent, *Nat. Inst. Med. Res.*, London).

Protein degradation also affects the Hedgehog signal-transduction pathway. The proteins Patched and Smoothed were thought to be components of a cell-surface receptor complex that is activated by Hedgehog protein to send signals inside the receiving cell. But it seems that this view may be simplistic. Although levels of *smoothed* messenger RNA are uniform across the epidermis and wing imaginal disc (the structure from which the wing will form), levels of Smoothed protein vary (S. Cohen, EMBL, Heidelberg; M. Noll, Univ. Zurich) (Fig. 1a, overleaf). This variation occurs because Patched promotes the degradation of Smoothed protein, except in cells that receive the Hedgehog signal. In these cells, Hedgehog-dependent removal of Patched from the cell surface results in increased phosphorylation of Smoothed and its accumulation at the cell membrane. These interactions complicate conventional models of receptor-mediated signalling.

So the protein-degradation machinery is required for normal regulatory processes, but it sometimes goes awry, with pathological consequences. Versions of the human protein Ataxin-A that have extra-long tracts of glutamine residues cause spinocerebellar ataxia type 1 (ref. 2). The expression of such human proteins in flies produces several defects seen in the human disease, including neuronal degeneration (J. Botas, Baylor Coll. Med., Houston). Mutations that disrupt components of the cellular 'ubiquitin-mediated' protein-degradation machinery suppress these defects. So, disease-causing Ataxin-A proteins may wreak their havoc by overloading the normal neuronal protein-degradation apparatus.

Protein modification, as well as turnover, is important in coordinating many signalling pathways, including the Notch pathway, discussed in these pages recently³. In places where the Notch protein establishes borders between groups of cells, its activity is regulated by Fringe (Fig. 1b, overleaf), an enzyme that adds particular sugar groups to Notch (and possibly to Notch's ligands; K. Irvine, Waksman Inst., Piscataway). A hint of further protein modification in Notch signalling comes from the localization of the Mastermind protein, a long-mysterious component of the pathway, to

Developmental biology

Post-expressionist flies

Sarah Bray and David Stein

The study of developmental biology is concerned with two main issues: the generation of cell diversity, and the organization of these diverse cells into beautifully patterned structures. These processes make use of relatively few, evolutionarily conserved cell-to-cell communication pathways. The mechanisms that coordinate the activities of these signalling pathways must be instrumental in fine-tuning the pattern of the developing organism. Over the past decade or so, emphasis has been placed on mechanisms of gene expression and its regulation as ways of coordinating cellular path-

ways. So it was exciting, at a recent meeting focusing on the fruitfly *Drosophila melanogaster*^{*}, to realize the variety of mechanisms that regulate the activity and subcellular localization of proteins in several signalling pathways.

Protein turnover is often neglected as a way of controlling cellular pathways, because most experimental methods detect steady-state levels of proteins in cells. But protein degradation appears to be important in the so-called Wingless and Hedgehog cell-to-cell signalling pathways. These pathways have several roles in the patterning of structures during *Drosophila* development. For example, they are involved in arranging the cuticle structures secreted by the embryonic

^{*}11th EMBO Workshop on Molecular and Developmental Biology of *Drosophila*, Greek Orthodox Academy, Kolymbari, Crete, 18–23 June 2000.

so-called PML bodies in the nucleus (S. Artavanis-Tsakonas, Mass. Gen. Hosp., Boston). These bodies are associated with 'SUMO modification'⁴ — a 'tagging' process that singles proteins out for degradation.

Signal transduction often involves protein modification by phosphorylation. Visualization of this process is helping to unravel the complexity of signalling from bone morphogenetic proteins (BMPs). A gradient of BMP proteins is thought to produce different cell fates both along the dorsal-ventral embryonic axis and across the wing⁵. The phosphorylation of Mad or Smad proteins is a key step in transmitting the BMP signal to the nucleus. Monitoring of Mad/Smad activation⁶ does reveal a gradient in the developing wing but, unexpectedly, there is no indication of graded activation of Mad/Smad along the embryonic axis (B. Shilo, Weizmann Inst.; T. Tabata, Univ. Tokyo; L. Raftery, Mass. Gen. Hosp., Boston). Instead, activation is detected only in a dorsal stripe of cells, whose fates require the highest levels of BMPs. This leaves unanswered the question of how signals from BMPs are propagated through the cells in which activation of Mad/Smad cannot be detected.

Signal-transduction pathways may also be influenced by the regulated targeting of their components to the nucleus. Nuclear localization of phosphorylated mitogen-activated protein kinase is the final step in an intracellular signalling pathway that begins

with the activation of a 'receptor tyrosine kinase' enzyme. The Corkscrew protein appears to control this final step by affecting the recruitment of importin (L. Perkins, Mass. Gen. Hosp., Boston), which carries proteins into the nucleus. A similar intersection with the nucleus-to-cytoplasm transport machinery affects the transcription factors Dorsal and Dif (C. Samakovlis, Umeå Univ., Sweden), which are involved in *Drosophila* immunity.

It is still too soon for the full impact of the recently completed *Drosophila* genome sequence to be felt. But one post-genomic revelation is the prevalence of gene families in *Drosophila*. For example, Warniu — a new member of the Snail protein family — is, like its two siblings, present in neural progenitor cells (T. Ip, Univ. Massachusetts). There are seven relatives of Wunen (R. Lehmann, Skirball Inst., New York), an enzyme that influences germ-cell migration⁷. A protein involved in ensuring that growing neuronal axons do not recross the embryonic midline is the receptor Roundabout⁸. Two more Roundabout-like proteins have now come to light, one of which also participates in the decision of axons to cross the midline (B. Dickson, Inst. Mol. Pathol., Vienna). Intriguingly, it appears that, after crossing the midline, the axons select their lateral position according to the combination of Roundabout proteins that they express.

The importance of regulated gene expression in development is undeniable. But our understanding of fly development must also incorporate changes in the stability, activity and localization of key proteins (and mRNAs too, as illustrated by their dramatic localization within the embryo; I. Davies, Univ. Edinburgh; D. Ish-Horowicz, Imperial Cancer Research Fund, London; H. Krause, Univ. Toronto). We now need to develop more techniques to look at both spatial (subcellular) and temporal mechanisms for coordinating protein activities and cell behaviours. The result will be an increasingly four-dimensional view of *Drosophila* development.

Sarah Bray is in the Department of Anatomy, University of Cambridge, Cambridge CB2 3DY, UK. e-mail: sjb32@mole.bio.cam.ac.uk

David Stein is in the Department of Molecular Genetics, Albert Einstein College of Medicine, New York, New York 10461, USA.

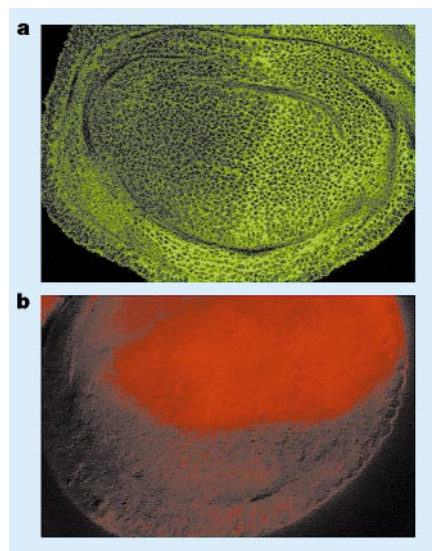


Figure 1 The importance of protein modification in *Drosophila* development. **a**, Levels of the Smoothed protein (green staining) vary across the wing imaginal disc. These levels are regulated by protein turnover⁹. **b**, *fringe* messenger RNA (red staining) is expressed more highly in the dorsal part of the wing disc. The Fringe protein adds particular sugar groups to the Notch receptor, modifying its interactions with its ligands. So this asymmetric distribution of *fringe* mRNA results in modification of Notch only in certain parts of the wing disc.

1. Sanson, B., Alexandre, C., Fascetti, N. & Vincent, J.-P. *Cell* **98**, 207–216 (1999).
2. Banfi, S. et al. *Nature Genet.* **7**, 513–520 (1994).
3. Fortini, M. E. *Nature* **406**, 357–358 (2000).
4. Müller, S., Matunis, M. J. & Dejean, A. *EMBO J.* **17**, 61–70 (1998).
5. Podos, S. D. & Ferguson, E. L. *Trends Genet.* **15**, 396–402 (1999).
6. Tanimoto, H., Itoh, S., ten Dijke, P. & Tabata, T. *Mol. Cell* **5**, 59–71 (2000).
7. Zhang, N., Zhang, J., Purcell, K. J., Cheng, Y. & Howard, K. *Nature* **385**, 64–67 (1997).
8. Kidd, T. et al. *Cell* **92**, 205–215 (1998).
9. Denef, N., Neubüser, D., Perez, L. & Cohen, S. M. *Cell* **102**, 521–531 (2000).

Daedalus

The sound of silence

The mobile phone, that essential modern accessory, makes its user immediately unpopular with those around him. For some reason, talking into such a phone is far more annoying to external listeners than a conversation with a human companion, or even soliloquial muttering. Daedalus reckons that the phone user instinctively projects his voice to reach the distant party. He is now inventing a phone which can be spoken into silently.

Speech is formed by the mouth and tongue acting as an ever-changing resonant cavity for tones produced by the larynx. The tones themselves are very basic; someone who has lost his larynx can speak intelligibly with a simple buzzer as a replacement. Daedalus's brilliant idea is to provide an ultrasonic 'buzzer' as a larynx. His 'Ultraphone' has a narrow pipe, like a drinking straw, which projects into the user's mouth and injects a set of inaudible ultrasonic frequencies into it. The user whispers or mouths his speech silently, and a microphone detects the modulations imposed by his mouth and palate on the ultrasonic signal. A heterodyne circuit downshifts this signal into the audio range, thus reconstituting the speaker's normal voice, and transmits it to the called party. Like a normal telephone, it also injects a proportion of the speaker's reconstituted speech back into his own earpiece as a 'side-tone' for aural feedback. Thus he hears his voice quite normally, and is not tempted to speak audibly. Indeed, any attempt to do so will result in strange distortions as the audio is downshifted and aliased by the heterodyne circuit.

This simple system would produce a flat and toneless speech, rather like that of the laryngectomy patient with his buzzer. But Daedalus hopes to equip the Ultraphone with a program that recognizes the tonal clues implicit in silently mouthed speech, and varies the ultrasonic frequencies in sympathy. This should give far more realistic speaking tones, close to the user's natural voice.

The Ultraphone will sweep the market. Yuppies and poseurs will be able to make truly silent phone calls anywhere, even in concert performances and prayer meetings, without disturbing the proceedings or revealing the important, confidential matters they are discussing. And even in a boiler factory or gunnery range, ambient noise will not distract them. High above the audible clamour, their ultrasonic deliberations will travel clear and unaccompanied.

David Jones

Obituary

Elvin A. Kabat (1914–2000)

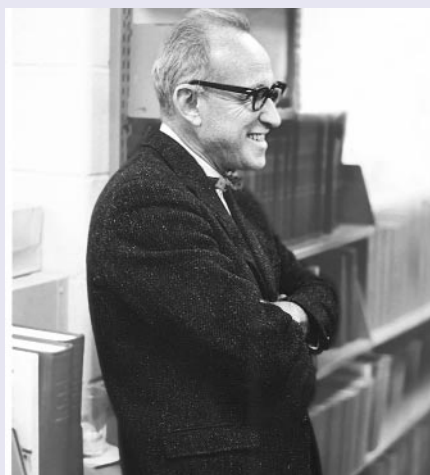
Elvin A. Kabat, who died on 16 June 2000, was a founder of modern immunochemistry. A man of great forthrightness and acerbic wit, he set the standard for immunological science for more than three decades.

Starting with his predoctoral work with Michael Heidelberger at Columbia University in New York, from 1933 to 1937, Kabat was a student of carbohydrate chemistry. He had a career-long interest in the structure and immunochemistry of blood-group substances, and determined the chemical basis of the ABO blood groups. This work, along with that of Walter Morgan and Winifred Watkins in Britain, laid the foundations for our current understanding of carbohydrate 'differentiation' antigens (cell-surface structures that vary according to developmental stage or tissue type) and their roles in cellular interactions.

Carbohydrate chemistry later allowed Kabat to analyse the binding properties of antibodies. Dextran — a polysaccharide found in yeast and bacteria, but not humans — was being used as a blood plasma substitute at the time. In 1951, after showing that dextran produces an immune response in humans, Kabat and his students started to use a series of oligosaccharides, containing different numbers of sugar subunits, to inhibit the interaction between dextran and the antibodies that recognize it. They found that the size and shape of the antibodies' dextran-binding sites varied; at most, they could accommodate six or seven single sugar units. This provided the first reliable estimate of the effective size of an antibody's antigen-binding site. These conclusions, as well as Kabat's later predictions that the shapes of antigen-binding sites can vary from shallow grooves to deep cavities, were eventually confirmed by X-ray crystallographers.

On moving to Uppsala in Sweden in 1937, Kabat worked for a year as a postdoctoral fellow with Arne Tiselius. While in Uppsala, Kabat provided the first physicochemical characterization of the antibodies known now as immunoglobulin G when he showed that they migrate electrophoretically in the gamma-globulin fraction of serum. Kabat's work led the way to all further studies of the chemical nature of immunoglobulins.

Kabat always had a keen interest in medicine. On returning to Columbia



Founder of modern immunochemistry

University in 1941, he developed an immunodiagnostic test for multiple sclerosis, and was among the first to develop an animal model of this autoimmune disease. These studies opened the way to modern experimental autoimmunity, with its promise of major benefit to patients with autoimmune diseases.

In 1970, Kabat recognized that the amino-acid sequences of immunoglobulins appearing in the literature provided an opportunity to determine the likely locations of their antigen-binding sites. He and Tai Te Wu developed the Wu–Kabat plot, which measures the distribution of variability along the immunoglobulin sequence. Their variability plots clearly identified the 'hypervariable' or 'complementarity-determining' regions, which vary the most from protein to protein, and the 'framework' regions, which vary the least. The hypervariable regions are those that contribute to the antigen-binding sites, and their variability means that different antibodies can recognize a huge range of antigens. From their plots, Wu and Kabat correctly predicted the locations of the antigen-binding regions in antibodies.

Following a year as a Fogarty scholar at the US National Institutes of Health (NIH) in 1974–75, Kabat divided his time between Columbia and the NIH for the rest of his career. We saw him often as he visited our laboratories or used our library. As protein and, later, DNA sequences became available, Kabat and his collaborators met at the NIH to collect together the amino-acid and DNA

sequences of immunoglobulins and other proteins of immunological importance. They produced a remarkable resource, *Sequences of Proteins of Immunological Interest*, which went through five printed editions. The book had an enormous impact on scientists who studied the relationships between the sequences and the antigen specificity of immunoglobulins, T-cell antigen receptors and molecules of the major histocompatibility complex. The Kabat database is now available on the Internet (<http://immuno.bme.nwu.edu/>), a living memorial to this great scientist.

Kabat received the National Medal of Science in 1991. He valued this honour greatly, particularly because of the difficulties he had in the 1950s when the NIH cravenly terminated his grants as a fallout of the politics of the McCarthy era. Fortunately, the Office of Naval Research and National Science Foundation continued to support him. Kabat saw the medal as recognition of a career-long record of accomplishment, and as a personal vindication.

But Kabat's influence went beyond his research. Among his published books was the landmark *Kabat and Mayer's Experimental Immunochemistry*, written with Manfred Mayer. This book was a bible for young immunologists in the 1960s. He also enthusiastically and uncompromisingly taught generations of medical and dental students in the laboratory and in lectures. His graduate students and postdoctoral fellows became leading lights in the immunological community. One, Baruj Benacerraf, became a Nobel laureate — an honour that eluded both Kabat and his mentor, Heidelberger, although many believed that their contributions amply merited this recognition.

Kabat was known for his devotion to science and for his view that scientists, as opposed to administrators, should determine how their field is organized. He was often the bane of university and government administrators in his complete commitment to science above all else. To fellow scientists, he was intellectually challenging, unsparing in his standards but always anxious to support them in their pursuit of scientific goals.

He is survived by his wife of 57 years, Sally (Linnick), three sons — Jon Kabat-Zinn, Geoffrey and David — and six grandchildren.

William E. Paul and Rose G. Mage
William E. Paul and Rose G. Mage are in the Laboratory of Immunology, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.
e-mail: wpaul@niaid.nih.gov

Membrane changes during hibernation

Organelle lipids undergo rapidly reversible rearrangement as body temperature drops.

Cellular membranes are susceptible to injury by cold, which causes their lipid components to separate^{1–4}. Here we investigate the structural changes that occur in organelle membranes inside cells of the central nervous system of hypothermic ground squirrels during hibernation. We find that lipids in these membranes sequester into protein-free domains that laterally displace membrane proteins and the underlying cytoplasmic matrix. But when the animal is aroused from hibernation, all these components return to their normal arrangements as the body temperature rises. Understanding this reversible temperature-induced redistribution of membrane lipids should help in the study of the effects of severe cold on non-hibernating species (including humans) and in the cryopreservation of cells and tissues.

Hibernation enables some mammals to survive cold temperatures and a shortage of food during the winter^{5,6}. The body of an animal preparing to hibernate drops almost to freezing temperature⁵ but warms up to 37 °C when it wakes at regular intervals for brief periods⁶. All organs, including the central nervous system (CNS), survive these rapid temperature shifts.

We used light and electron microscopy to analyse the CNS and other tissues during hibernation cycles of adult ground squirrels (*Spermophilus tridecemlineatus*)⁷. When the temperature fell in these hibernating animals, neuronal and glial cells from a variety of nervous-system tissues, including the hippocampus and trigeminal ganglia, underwent widespread structural changes that manifested in the form of visible 'slits' in the cytoplasm (Fig. 1). The size of these slits ranges up to 10 µm in length and up to 2 µm in width; they were not detected in non-hibernating squirrels or in euthermic squirrels four hours after arousal from hibernation.

The slits (Fig. 1c) correspond to segments of endoplasmic reticulum (ER) membranes that do not carry ribosomes and lack underlying cisternal and cytoplasmic matrix. These segments were contiguous or adjacent to ER membranes with normal morphology. We observed similar structural changes in several other neuronal and non-neuronal tissues, including the organ of Corti, the cochlear stria vascularis, the pituitary and the adrenal glands.

We also analysed these hibernation-induced changes by freeze-fracture electron microscopy. In tissues from normothermic, awake squirrels, the fracture faces of the ER of all cells showed the typical random dis-

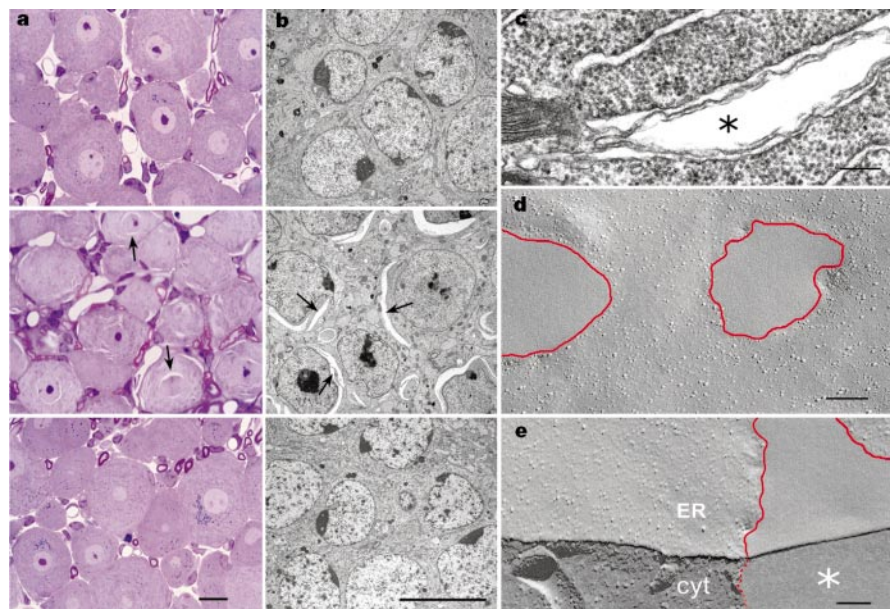


Figure 1 Structural changes induced by low temperature in endoplasmic reticulum membranes during hibernation. **a**, Light- and **b**, electron-microscopy views of trigeminal ganglion (**a**) and hippocampal (**b**) neurons before (top), during (middle), and four hours after (bottom) deliberate arousal from hibernation. Cytoplasmic 'slits' (arrows) are widespread during hibernation only when the body temperature was about 4 °C. **c**, Higher magnification reveals that the slits correspond to segments of endoplasmic reticulum (ER) cisternae that are devoid of ribosomes and underlying cytoplasmic matrix (asterisk). **d,e**, Freeze-fracture images of ER membranes from a hippocampal neuron (**d**) and an epithelial cell of the stria vascularis (**e**) show protein-free patches (outlined in red). The cross-fracture view of the cytoplasm (cyt) shows that the cytoplasmic matrix is cleared from the region underlying the protein-free domain (asterisk). Scale bars, 10 µm in **a,b**; 0.2 µm in **c–e**.

tribution of intra-membrane protein particles. In tissues from hypothermic, hibernating squirrels, the ER membranes showed areas devoid of proteins, generally seen as circular patches (Fig. 1d).

The size of the protein-free domains varied from small circular 0.2–2-µm patches in hippocampal and spinal-cord neurons and in cerebellar Purkinje cells, to over 10-µm-wide areas in pituitary cells and in hippocampal neurons. The relative distribution in size and frequency of the protein-free patches appeared very similar to the slit-shaped clear cytoplasm associated with ER membranes seen in thin sections (assuming that the slits represent cross-sections of the protein-free membrane patches). We confirmed these results using unfixed, flash-frozen tissues.

The cross-fractured cytoplasm underlying the protein-free patches also had a smooth background, indicating that the ribosomes and the cytoplasmic matrix had been displaced. In some CNS neurons and astrocytes, small protein-free areas were noted in the plasma membrane (not shown). No protein-free areas were detected in the ER or in the plasma membranes in the pancreas, liver or intestine.

When the temperature is lowered in

some membranes *in vitro*, the 'gelation' of high-melting-point lipids induces lipid segregation and the lateral displacement and exclusion of proteins^{1–3}. The protein-free domains we observe in hibernating animals may result from this intrinsic phase behaviour of membrane lipids, which may participate in the cellular mechanism of low-temperature acclimation.

Previous studies of cold adaptation during hibernation have focused largely on seasonal changes in the lipid composition of tissues, which reflect a stepwise preparation for hibernation^{8,9}. Such slow alterations cannot explain the sudden acclimation to rapid temperature shifts during arousal and re-entry into hibernation. As our results reveal, changes in lipid composition do not prevent the temperature-induced changes in membrane structure we observe in several tissues of the ground squirrel, including the CNS. On the contrary, they may permit reversibility of these membrane-structure changes and counteract their potentially deleterious effects.

Nabil A. Azzam*, **John M. Hallenbeck***, **Bechara Kachar†**

*Stroke Branch, National Institute of Neurological Disorders and Stroke, and †Section on Structural Cell Biology, National Institute on Deafness and

other Communication Disorders, National Institutes of Health, Bethesda, Maryland 20892, USA
e-mail: kacharb@nidcd.nih.gov

1. Arav, A. et al. *Cryobiology* **33**, 589–599 (1996).
2. Crowe, J. H. et al. *Cryobiology* **38**, 180–191 (1999).
3. Williams, W. P. *Phil. Trans. R. Soc. Lond. B* **326**, 555–570 (1990).
4. Wada, H. et al. *Nature* **347**, 200–203 (1990).

Metabolism

Winter torpor in a large bird

Torpor is a natural state in which animals show a substantial and controlled reduction of body temperature to conserve energy^{1,2}. A few small birds (weighing less than 80 g) are known to use it as a survival strategy in winter, but we have discovered that a large bird, the Australian tawny frogmouth, which weighs 500 g, can also enter this state. This surprising finding increases the size of birds known to use natural torpor by almost tenfold, suggesting that avian torpor is more widespread than is commonly believed, enabling birds to stay in their territory throughout the year.

Small endothermic birds and mammals have enormous food requirements because of the fast metabolic rate that is necessary to regulate a high body temperature (T_b). When this high metabolic rate is unsustainable, for example in periods of adverse weather and/or food shortage, many small mammals (body mass 2–10,000 g) survive by entering a state of torpor^{1–3}. The study of torpor in birds has so far been restricted to species weighing less than 80 g (refs 2,4,5), and rather than using torpor to overcome

5. Barnes, B. M. *Science* **244**, 1593–1595 (1989).
6. Lyman, C. P., Willis, J. S., Malan, A. & Wang, L. C. H. *Hibernation and Torpor in Mammals and Birds* (Academic, New York, 1982).
7. Frerichs, K. U. et al. *Proc. Natl Acad. Sci. USA* **95**, 14511–14516 (1998).
8. Aloia, R. C. *Fed. Proc.* **39**, 2974–2979 (1980).
9. Geiser, F., Kenagy, G. J. & Wingfield, J. C. *J. Comp. Physiol. B* **167**, 416–422 (1997).

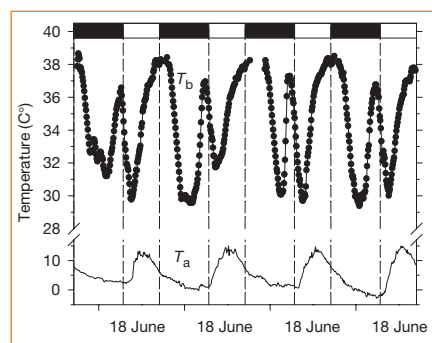


Figure 2 Fluctuations in body temperature of large Australian tawny frogmouth (T_b , filled symbols, measured by an internal transmitter) and in air temperature (T_a , solid line) over 4 days in June (winter). T_b was high at the beginning of the night (black bars), and then declined. The bird was aroused before dawn, after which it changed its roost and re-entered torpor. T_b increased from mid-morning to peak around sunset.

adverse conditions, birds may migrate to avoid them. However, many birds are sedentary and often rely on ephemeral, weather-dependent food sources — so how do they overcome periodic energy bottlenecks?

To answer this question, we investigated whether the Australian tawny frogmouth (*Podargus strigoides*; Fig. 1), a sedentary bird which feeds mainly on arthropods, uses torpor in the wild. The study was conducted from the Australian autumn to summer in an open woodland of *Eucalyptus* and *Acacia* at 1,000 m altitude in a cool temperate area near Armidale, New South Wales. We captured seven frogmouths and fitted them with temperature-sensitive transmitters (calibrated to the nearest 0.1 °C) weighing 3 g. All birds received an external backpack-style transmitter⁴ (long range) to measure skin temperature (T_{skin}), and three birds had a second internal transmitter (short range), to measure core T_b and to determine $T_b - T_{skin}$ differentials, implanted under general anaesthesia. Transmitter signals were recorded at 10-min intervals for up to nine months⁶.

All individuals entered torpor in winter: T_b fluctuated around 38–40 °C during activity and fell to about 36 °C during the rest phase, with a lower limit of 34 °C. On cold (less than 7 °C) winter nights in June–August, T_b fell below 34 °C on 202 of 462 observation days (44%). The minimum T_b recorded by an internal transmitter was 29.1 °C, the lowest T_b calculated from T_{skin} and the $T_b - T_{skin}$ differential was 27.2 °C, and the mean minimum T_b for the seven birds was 29.0 ± 1.0 °C. Birds usually

entered torpor after a brief phase of activity in the evening (Fig. 2), became active again near sunrise and frequently entered a second bout of torpor after they had flown to their day roost. Dawn torpor was briefer (3.5 ± 1.2 hours) than night torpor (7.0 ± 1.2 hours) and was often terminated by passive rewarming in the sun.

Our study shows that the large frogmouth regularly enters torpor in winter. We argue that the use of torpor enables the bird to survive in a wide range of habitats and to remain resident in its territory throughout the year⁷, despite feeding on arthropods. Because torpor occurs in this large bird and because birds are more diverse and smaller on average than mammals, we predict that avian torpor is much more common than is currently believed.

Gerhard Körtner, R. Mark Brigham*, Fritz Geiser

School of Biological Sciences, University of New England, Armidale 2351, Australia
e-mail: fgeiser@metz.une.edu.au

*Permanent address: Department of Biology, University of Regina, Regina, Saskatchewan S4S 0A2, Canada

1. Lyman, C. P., Willis, J. S., Malan, A. & Wang, L. C. H. *Hibernation and Torpor in Mammals and Birds* (Academic, New York, 1982).
2. Geiser, F. & Ruf, T. *Physiol. Zool.* **68**, 935–966 (1995).
3. Cossins, A. R. & Barnes, B. M. *Nature* **382**, 582–583 (1996).
4. Brigham, R. M. *Physiol. Zool.* **65**, 457–472 (1992).
5. Reintertsen, R. E. *Polar Res.* **1**, 269–284 (1983).
6. Körtner, G. & Geiser, F. *Oecologia* **123**, 350–357 (2000).
7. Körtner, G. & Geiser, F. *J. Zool. Lond.* **248**, 501–507 (1999).

Ageing

Cloning of mice to six generations

Mice have been cloned by nuclear transfer into enucleated oocytes^{1–3}, and here we describe the reiterative cloning of mice to four and six generations in two independent lines. Successive generations showed no signs of premature ageing, as judged by gross behavioural parameters, and there was no evidence of shortening of telomeres at the ends of chromosomes, normally an indicator of cellular senescence — in fact, these appeared to increase slightly in length. This increase is surprising, given that the number of mitotic divisions greatly exceeds that of sexually produced animals and that any deleterious effects of cloning might be expected to be amplified in sequentially cloned mice. Our results offer a new approach to the study of organismal ageing.

Founder clones (G1) of two lines, A and B, were generated using cumulus cells of adult female B6C3F1 mice (agouti coat colour) as nucleus donors, and oocytes from adult B6D2F1 (black) females as nucleus recipients. The cloning procedure



Figure 1 Tawny frogmouth (*Podargus strigoides*).

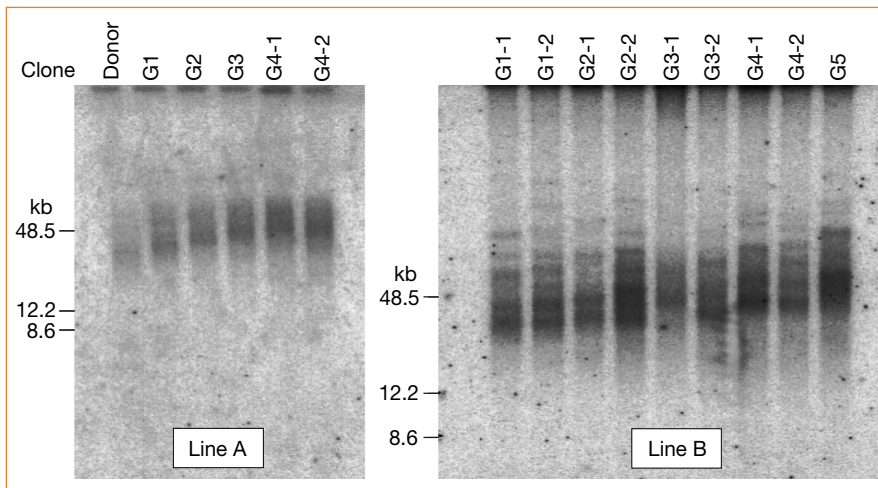


Figure 1 Telomere lengths in successive generations (G1–G5) of mice cloned from cumulus cells. Southern-blot analysis of terminal restriction-enzyme-cut fragments in five sequential generations shows that telomeres do not undergo incremental erosion in successive clonal generations. Genomic DNA isolated from peripheral-blood lymphocytes taken from representative animals from each generation was digested with the restriction enzyme *HinfI*, resolved on a pulse field gel, transferred to a solid support and probed with a 5'-³²P-labelled (T₄AG)₃ oligonucleotide. Peripheral blood lymphocytes were sampled on the same day. Ages of mice (in months) were: in line A, donor, 18; G1, 16; G2, 14; G3, 12; G4, 9; G5, 9; in line B, G1, 15.5; G2, 13; G3, 11; G4, 9; G5, 7. Suffix numbers (G4-1 and G4-2, for example) identify different pups of each generation.

was repeated with cumulus cells from adult G1 mice as nucleus donors to produce the next clonal generation, G2, and so on. Table 1 summarizes the results obtained following the reconstruction of 3,920 enucleated oocytes.

Previously, about 2% of enucleated oocytes receiving a cumulus cell nucleus developed to live-born pups¹. This value is comparable to the cloning efficiency for G1 in lines A (1.5%) and B (4.2%). However, the success rate dropped in successive cloned generations: line A did not produce a G5 clone from 670 reconstructed oocytes; in line B, the only live-born G6 clone was cannibalized by her foster mother, thereby terminating the line. Mouse lines A and B therefore represent totals of 9 and 26 clones from their respective donors. Placental size was consistently in the range previously reported for cloned mice² and did not increase in successive generations (data not shown).

Do sequentially cloned mice show signs of accelerated ageing? We assessed the behaviour of these mice and determined telomere lengths to assess organismal and cellular measures of ageing, respectively. We evaluated learning ability in the Morris water maze and Krushinsky tests, as well as strength and agility, and also used other

assays designed to monitor signs of pre-mature ageing, such as a decline in activity in the home cage and loss of coordination⁴. All cloned mice were, by these criteria, normal compared with age-matched controls (data not shown); the G5 mouse is alive and healthy at 1.5 years.

We measured telomere length in peripheral blood lymphocytes of clones G1–G6 by Southern-blot analysis of terminal restriction-enzyme-digested fragments (Fig. 1) and found no evidence of shortened telomeres in the cloned mice. In fact, our results show that the telomeres lengthen with each generation. As representative animals of each generation were sampled simultaneously, we cannot rule out an age-related contribution to this increase (with younger mice having longer telomeres). In addition, long telomeres in mice are optimally studied by means of different assays such as quantitative fluorescence *in situ* hybridization⁵. We have detected telomerase activity in cumulus cells (data not shown); it is therefore possible that telomeres in these cells are unusually long, resulting in offspring with concomitantly longer telomeres.

Shortened⁶ and lengthened⁷ telomeres have been reported in cloned livestock but, unlike ours, those experiments involved only a single round of cloning. Our results

on sequentially cloned mice verify that telomere shortening is not a necessary outcome of the cloning process⁸. However, as only 1–2% of reconstructed oocytes yield live-born clones, the possibility of selection for donor nuclei with the longest telomeres cannot be excluded. Further investigation is required into the consequences of nuclear transfer on telomere length and lifespan.

Teruhiko Wakayama*†, Yoichi Shinkai‡, Kellie L. K. Tamashiro*, Hiroyuki Niida§, D. Caroline Blanchard||, Robert J. Blanchard||, Atsuo Ogura¶, Kentaro Tanemura¶, Makoto Tachibana‡, Anthony C. F. Perry#, Diana F. Colgan#, Peter Mombaerts#, Ryuzo Yanagimachi*

*Department of Anatomy and Reproductive Biology, John A. Burns School of Medicine, University of Hawaii, Honolulu, Hawaii 96822, USA

‡Department of Cell Biology, Institute for Virus Research, Kyoto University, Sakyo-Ku, Kyoto 606-8507, Japan

§Life Sciences Division, Lawrence Berkeley National Laboratory, University of California, Berkeley, California 94720, USA

||Bekesy Laboratory of Neurobiology, University of Hawaii, Honolulu, Hawaii 96822, USA

¶Department of Veterinary Science, National Institute of Infectious Disease, Tokyo 162-8640, Japan

#The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA

†Present address: The Rockefeller University, 1230 York Avenue, New York, New York 10021, USA
e-mail: wakayata@mail.rockefeller.edu

1. Wakayama, T., Perry, A. C. F., Zuccotti, M., Johnson, K. R. & Yanagimachi, R. *Nature* **394**, 369–374 (1998).
2. Wakayama, T. & Yanagimachi, R. *Nature Genet.* **22**, 127–128 (1999).
3. Wakayama, T., Rodriguez, I., Perry, A. C. F., Yanagimachi, R. & Mombaerts, P. *Proc. Natl Acad. Sci. USA* **96**, 14984–14989 (1999).
4. Tamashiro, K. L., Wakayama, T., Blanchard, R. J., Blanchard, D. C. & Yanagimachi, R. *Biol. Reprod.* **63**, 328–334 (2000).
5. Zijlmans, J. M. et al. *Proc. Natl Acad. Sci. USA* **94**, 7423–7428 (1997).
6. Shiels, P. G. et al. *Nature* **399**, 316–317 (1999).
7. Lanza, R. P. et al. *Science* **288**, 665–669 (2000).
8. Wilmut, I., Clark, J. & Harley, C. B. *Nature Biotech.* **18**, 599–600 (2000).

Gene expression

Total silencing by intron-spliced hairpin RNAs

Post-transcriptional gene silencing (PTGS), a sequence-specific RNA degradation mechanism inherent in many life-forms, can be induced in plants by transforming them with either antisense¹ or co-suppression² constructs, but typically this results in only a small proportion of silenced individuals. Here we show that gene constructs encoding intron-spliced RNA with a hairpin structure can induce PTGS with almost 100% efficiency when directed against viruses or endogenous

Table 1 Effect of sequential cloning on full-term development

Line	G1	G2	G3	G4	G5	G6	Total
A	2/131 (1.5)	1/228 (0.4)	1/263 (0.4)	5/238 (2.1)	0/670 (0)	-	9/1,530 (0.6)
B	4/96 (4.2)	7/351 (2.0)	5/352 (1.4)	6/286 (2.1)	3/581 (0.5)	1/724 (0.1)	26/2,390 (1.1)

Successive generations are represented as G1, G2 and so on for two independent mouse lines, A and B. The number of pups born live after cumulus-cell nuclear transfer is compared to successfully reconstructed oocytes (pups/oocytes), with the corresponding percentages in parentheses. Significant χ^2 comparisons were derived for G4 and G5 from line A, G1 and G5, G6 from line B, and G2, G3, G4 versus G6 from line B ($P < 0.05$).

genes. These constructs could prove valuable in reverse genetics, genomics, engineering of metabolic pathways and protection against pathogens.

Induction of PTGS by co-suppression and antisense methods that target the *Nia*-protease (*Pro*) gene sequence of potato virus Y (PVY)³ cause silencing in 7% and 4% of independent transformants, respectively; induction of PTGS in these tobacco plants (*Nicotiana tabacum*) manifests as immunity^{3,4} to the virus.

Using principles we developed for silencing constructs that express double-stranded RNA and inverted-repeat RNAs³, we made a construct encoding a single self-complementary hairpin RNA (hpRNA) of the *Pro* sequence. The construct contains sense and antisense *Pro* sequences flanking an 800-nucleotide spacer fragment derived from the *uidA* (GUS) gene (Fig. 1a). About 60% (25/43) of the plants that are transformed with this construct, many of which contained a single transgene copy, were immune to the virus. The spacer fragment contributed to the stability of the perfect inverted-repeat sequences, but it was not required for the specificity of the PTGS (Fig. 1a).

To test the effect of removing the loop region of hpRNA, we replaced the spacer with an intron sequence (Fig. 1a, b). The intron sequence provides stability to the DNA, but is spliced out during pre-mRNA processing⁵ to produce loopless hpRNA. As a control, we made a sister construct in which the intron sequence was inserted in the reverse, non-splicing, orientation. When transformed into tobacco, 22 of 34 (65%) reverse-intron plants were immune, a similar frequency to plants transformed with the GUS spacer construct. Amazingly, we found that 22 of 23 plants transformed with the construct containing the functional intron were immune to the virus.

To test whether this enhancement by the sense-intron construct was a general phenomenon, we made two hpRNA constructs against the endogenous $\Delta 12$ -desaturase (*Fad2*) gene of *Arabidopsis*, which catalyses the conversion of oleic to linoleic acid in the seed^{6,7}; one construct contained an intron and the other a non-intron spacer region. We found that 69% (44/63) of the transgenic plants with the non-intron spacer region construct showed PTGS of the $\Delta 12$ -desaturase gene, but that 100% (30/30) of plants transformed with the intron construct showed silencing of this gene.

How does the presence of this intron enhance silencing efficiency? The process of intron excision from the construct by the spliceosome might help to align the complementary arms of the hairpin in an environment favouring RNA hybridization, promoting the formation of a duplex. Alternatively, splicing may transiently increase

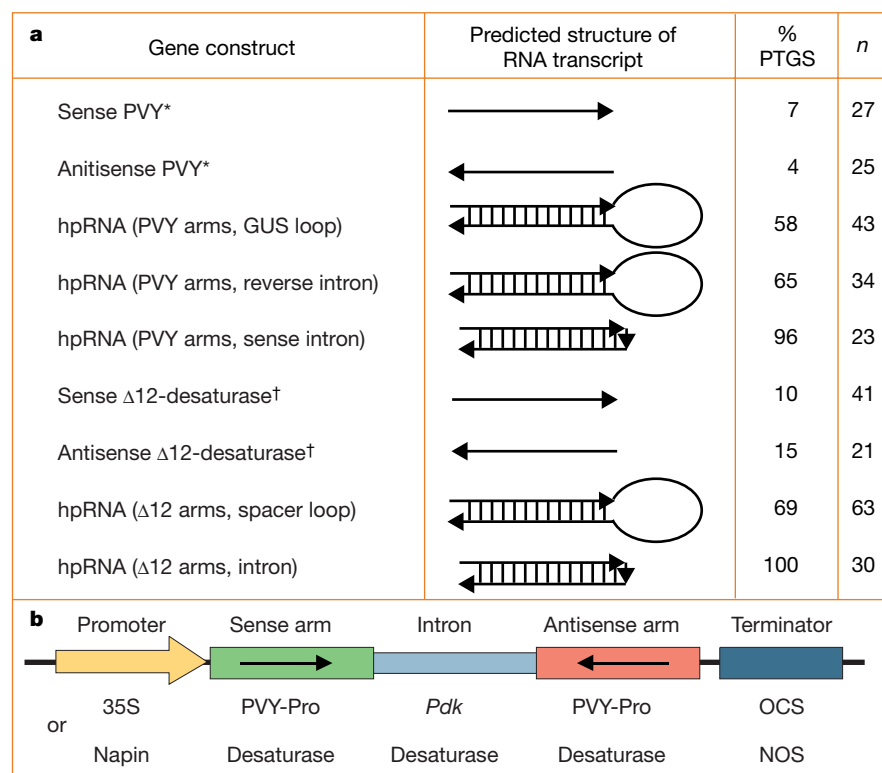


Figure 1 Efficiency of induction of post-transcriptional gene silencing (PTGS) by different gene constructs and the predicted structure of RNA transcribed from the transgenes. **a**, PTGS efficiency measured for potato virus (PVY) and $\Delta 12$ -desaturase genes as the percentage of independent transgenic plants immune to PVY and the percentage of plants with enzyme activity reduced by more than 20% compared with wild type, respectively. In the predicted structures of RNA transcripts, right- and left-pointing arrows represent sense and antisense orientation of sequences, respectively; small vertical arrows represent splice-junction sequences remaining after the intron has been spliced out. Vertical lines in the predicted structures indicate duplex formation. Asterisks, data from ref. 3; daggers, data from ref. 7; hpRNA, hairpin RNA; n, number of independent transformants; GUS, β -glucuronidase. **b**, Design of intron-containing hairpin constructs. OCS, octopine synthase; NOS, nopaline synthase.

the amount of hairpin RNA by facilitating, or retarding, the hairpin's passage from the nucleus, or by creating a smaller, less nuclease-sensitive loop.

Our PVY constructs contained intron-2 from the *Pdk* gene of *Flaveria*⁸, whereas the $\Delta 12$ -desaturase construct contained intron-1 from the *Arabidopsis Fad2* gene (Fig. 1b). PVY constructs were controlled by the constitutive CaMV35S (ref. 9) promoter and produced hpRNA containing the PVY coding-region sequence (700 nucleotides); the desaturase gene construct used the seed-specific napin promoter¹⁰ to produce hpRNA representing 120 nucleotides of the 3'-untranslated region of the $\Delta 12$ -desaturase gene.

We believe that constructs encoding intron-hpRNA should efficiently induce PTGS for a wide range of genes in a variety of circumstances and could become as useful to plant biology as RNAi^{11,12} is to the study of nematodes and *Drosophila*. The transgene design might also have applica-

tion in organisms in which RNAi has been obtained by injection of double-stranded RNA.

Neil A. Smith, Surinder P. Singh, Ming-Bo Wang, Peter A. Stoutjesdijk, Allan G. Green, Peter M. Waterhouse
CSIRO Plant Industry, Canberra, ACT 2601, Australia
e-mail: peter.waterhouse@pi.csiro.au

- Hamilton, A. J., Lycett, G. W. & Grierson, D. *Nature* **346**, 284–287 (1990).
- Jorgensen, R. A. *Science* **268**, 689–691 (1995).
- Waterhouse, P. M., Graham, M. W. & Wang, M.-B. *Proc. Natl. Acad. Sci. USA* **95**, 13959–13964 (1998).
- Waterhouse, P. M., Smith, N. A. & Wang, M.-B. *Trends Plant Sci.* **4**, 452–457 (1999).
- Wang, M.-B., Upadhyaya, M. N., Brettell, R. I. S. & Waterhouse, P. M. *J. Genet. Breed.* **51**, 325–334 (1997).
- Okuley, J. *et al. Plant Cell* **6**, 147–158 (1994).
- Cartea, M. E., Migdal, M., Galle, A. M., Pelletier, G. & Guerche, P. *Plant Sci.* **136**, 181–194 (1998).
- Rosche, E. & Westhoff, P. *Plant Mol. Biol.* **29**, 663–678 (1995).
- Gleave, A. P. *Plant Mol. Biol.* **20**, 1203–1207 (1992).
- Stalberg, K., Ellerstrom, M., Josefsson, L. G. & Rask, L. *Plant Mol. Biol.* **23**, 671–683 (1993).
- Fire, A. *et al. Nature* **391**, 806–811 (1998).
- Sharp, P. A. *Genes Dev.* **13**, 139–141 (1999).

brieF communications is intended to provide a forum for short, topical reports of general scientific interest and for technical discussion of recently published material of particular interest to non-specialist readers. Priority will be given to contributions that have fewer than 500 words, 10 references and only one figure. Detailed guidelines are available on *Nature's* website (www.nature.com) or on request from nature@nature.com

Loosening of plant cell walls by expansins

Daniel J. Cosgrove

Department of Biology, 208 Mueller Laboratory, Pennsylvania State University, University Park, Pennsylvania 16802, USA

Plant cell walls are the starting materials for many commercial products, from lumber, paper and textiles to thickeners, films and explosives. The cell wall is secreted by each cell in the plant body, forming a thin fibreglass-like network with remarkable strength and flexibility. During growth, plant cells secrete a protein called expansin, which unlocks the network of wall polysaccharides, permitting turgor-driven cell enlargement. Germinating grass pollen also secretes an unusual expansin that loosens maternal cell walls to aid penetration of the stigma by the pollen tube. Expansin's action has puzzling implications for plant cell-wall structure. The recent explosion of gene sequences and expression data has given new hints of additional biological functions for expansins.

For plants, the cell wall is an important structure that determines cell shape, glues cells together, provides essential mechanical strength and rigidity, and acts as a critical barrier against pathogens. Secreted by growing cells, the primary wall is a polymeric network of crystalline cellulose microfibrils embedded in a hydrophilic matrix of hemicelluloses and pectins¹ (Box 1). Unlike the bacterial wall, which may form one giant, covalently linked macromolecule, the polysaccharides of the growing plant cell wall are mostly separate long-chained polymers that form a cohesive network through non-covalent lateral associations and physical entanglements.

The growing wall possesses a remarkable combination of strength and pliancy, enabling it to withstand the large mechanical forces that

arise from cell turgor pressure, while at the same time permitting a controlled polymer 'creep' that distends the wall and creates space for the enlarging protoplast. With a tensile modulus greater than 10^{11} N m^{-2} , cellulose microfibrils themselves are effectively inextensible. Wall expansion occurs by slippage or rearrangement of the matrix polymers that coat the microfibrils and hold them in place. Until recently this was thought to occur primarily by hydrolysis of matrix polysaccharides, but the discovery of expansins has uncovered another mechanism of wall enlargement.

Discovery of expansins

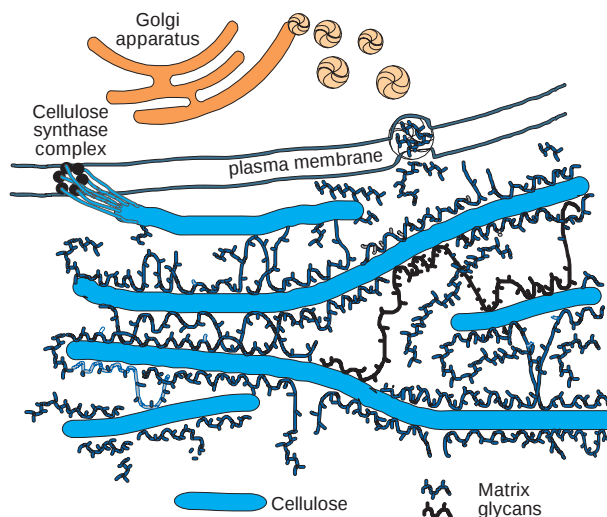
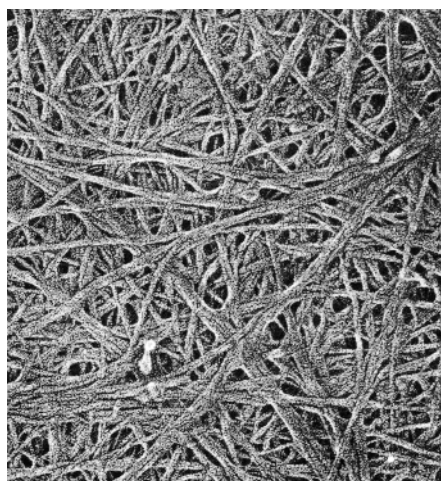
Plant cells originate mainly in small pockets of dividing cells, called meristems, that are located at the growing points of shoots, roots and

Box 1

Structure and biogenesis of the primary cell wall in plants

All land plants make a primary cell wall that is remarkably similar in structure to, but distinct from, that of most algae, fungi, bacteria and other groups with cell walls. The left-hand panel of the Figure shows the microfibrillar nature of the wall as a 'relief map' of the cell-wall surface after extraction of most of the matrix polysaccharides (image from ref. 4, used with permission). The right-hand panel is a diagram of cell-wall assembly. Cellulose microfibrils are synthesized by large complexes in the plasma membrane. Newly secreted cellulose then associates with matrix glycans (hemicelluloses and pectins) which are synthesized in the Golgi apparatus and delivered to the wall by

secretory vesicles. The cellulose microfibril is a thin ribbon, about 5 nm in width and many micrometers in length. It consists of an ordered array of many parallel chains of an unbranched glucose polymer (1,4- β -glucan). Hemicelluloses typically are branched polysaccharides characterized by a strong tendency to bind to cellulose, whereas pectins are generally acidic polysaccharides with a strong tendency to form ionic gels. A small amount of structural protein is also found on the wall, but its function is uncertain.



other organs. Meristematic cells are small (~5 μm) and densely packed with cytoplasm. When cells are displaced from the meristem, they typically undergo a prolonged phase of enlargement and differentiation during which cell volume greatly increases. To take an extreme example, a water-conducting cell of the xylem may be a million times larger in volume than its meristematic initial. Such cell enlargement is achieved economically by filling the cell with a large central vacuole containing water and solutes. The physical control of this process resides in the ability of the cell wall to undergo turgor-driven expansion⁸.

Growing plant cell walls typically extend faster at low pH, a phenomenon referred to as 'acid growth'⁹. This is readily seen in isolated walls clamped at constant tension in an extensometer (Fig. 1a). At neutral pH the walls soon stop extending but they rapidly extend when pH is lowered (Fig. 1b). Acid-induced extension is not merely a physical property of wall polysaccharides, but requires active wall proteins¹⁰. A key breakthrough¹¹ came with the discovery that wall proteins may be added back to denatured cell walls to restore their ability to extend (Fig. 1c). The active proteins were of a new protein class, subsequently named expansin¹². Their action is referred to as 'wall loosening', a general term that does not specify the exact mechanism for induction of cell wall extension.

Expansins, proteins with relative molecular mass of about 26,000, were isolated first from young cucumber seedlings and subsequently from other plant tissues, where their activity was associated with cell growth. Expansins are also implicated in the drought responses of maize seedlings, where maintenance of root growth involves increased expansin activity in the growing region¹⁵. Recent work indicates that drought increases the expression of expansin genes in a spatial and temporal pattern that closely matches the changes in expansin protein activity (Y. Wu, R. E. Sharp and D.J.C., manuscript in preparation). It may thus be possible to engineer enhanced drought tolerance into crop plants by manipulating expansin gene expression.

Other studies have shown close correlations between expansin gene expression and growth. For example, deep-water rice responds to submergence with rapid stem elongation, which is preceded by large increases in expansin messenger RNA¹⁶. The growth hormone gibberellin has similar effects on stem elongation and expansin

expression. In the wetland plant *Rumex palustris*, submergence stimulates elongation and expansin expression, but this stimulation is not found in a related species that lacks the submergence growth response¹⁷. Auxin-induced root formation in pine cuttings is accompanied by a 100-fold increase in expansin mRNA¹⁸. Finally, in the shoot apical meristem, which produces leaf primordia in a carefully orchestrated spatial and temporal pattern, an expansin gene is expressed at the site of future primordium emergence, preceding the first morphological signs of primordium outgrowth¹⁹. These results confirm the connection between expansin gene expression and growth induction.

Consistent with these gene-expression results, other studies found that cells grew faster upon treatment with expansin proteins^{20,21}. Moreover, when expansin-loaded microbeads were applied locally to shoot apical meristems at the sites of incipient leaf primordia, the primordia emerged prematurely²². The resulting out-of-sequence leaf emergence caused the spiral pattern of primordium initiation (phyllotaxy) on the meristem dome to reverse²³.

The concept that expansins are key endogenous regulators of plant cell enlargement is now supported by four lines of evidence. First, expansin proteins induce stress relaxation and extension of isolated cell walls in a pH-dependent manner^{11,24}. No other proteins have been found with this activity^{7,25}. Second, application of expansins to living cells stimulates cell enlargement. This shows that expansins can function under the normal conditions found in the cell walls of living cells and that they are present in sub-saturating amounts in these tissues. Third, expansin genes are expressed at the right time and in the right place to function in growth control. And finally, reduction of expansin gene expression by antisense methods inhibits growth³². Thus, the concept of expansins as catalysts of cell-wall enlargement is well supported.

Allergens, β -expansins and grass cell walls

When the first expansin complementary DNAs were sequenced³³, BLAST searches of GenBank revealed a distant sequence similarity to a group of grass pollen allergens called group-1 allergens³⁴. These pollen proteins were identified by immunologists in the mid-1980s as the main causative agents of hay fever and seasonal asthma induced by grass pollen³⁵. Assays of the group-1 allergen from maize

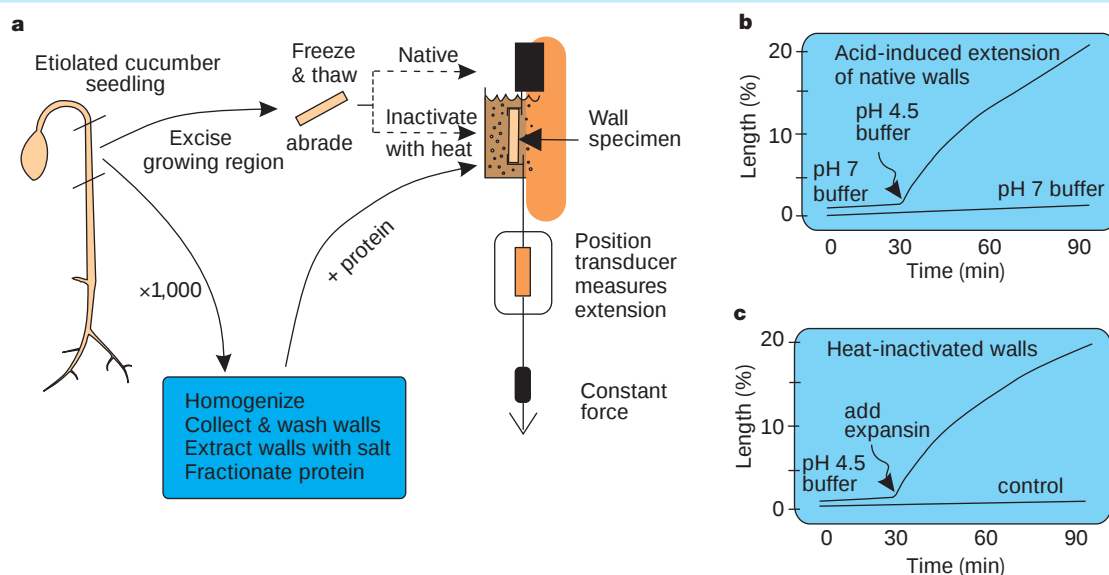


Figure 1 Diagram of extensometer assays. **a**, The growing hypocotyl of a seedling is cut and frozen to kill the cells. The wall specimen is either directly clamped in a constant-force extensometer ('native walls') or first inactivated with a brief heat treatment before being clamped in the extensometer ('heat-inactivated walls'). Expansin protein is prepared by extraction from native walls, followed by fractionation and addition to the wall. **b**, Native walls extend very little at neutral pH, but rapidly extend in acidic pH. **c**, Heat-inactivated walls lack acid-induced extension, which can be restored by addition of expansin to the walls.

pollen (called Zea m1) showed it to have wall-loosening activity characteristic of expansin³⁶, but with an intriguing twist: the allergen induced extension only in grass cell walls and was not effective on walls from dicotyledons. This activity has since been confirmed for group-1 allergens from other grasses (L.-C. Li and D.J.C., unpublished data). This discovery has several implications.

First, it appears that group-1 allergens have a wall-loosening role, aiding penetration of the pollen tube through the stigma and style by softening the maternal cell walls. This function has not been proven, but is indicated by the unusually high abundance and solubility of these proteins³⁶. These properties would be anomalous if the group-1 allergens were targeted for action on the growing pollen-tube wall, which consists of a very small area at the tip of the pollen tube. However, they make sense if the target were the many walls between which the pollen tube must push on its way to the ovary³⁷. Indeed, experimental measurements³⁶ show that maize style walls are readily loosened by Zea m1.

Second, it appears that grasses are the only flowering plants to have discovered the 'trick' of using a pollen-secreted expansin to assist pollen-tube penetration. This inference is based on the finding that pollen extracts from other plant groups lack expansin-like loosening activity³⁶. Moreover, BLAST searches of GenBank show homologous pollen allergens only in grasses. If grasses have discovered this method of invading plant tissues, however, can fungi and bacteria be far behind? These organisms might possess novel expansins.

A third inference is that the group-1 allergens act on polymers specific to grass walls, which differ from the typical plant cell wall in several regards. In particular, grasses have reduced the usual matrix polysaccharides (xyloglucan and pectin) and substituted other polysaccharides, most notably branched xylans and mixed-linked glucan³⁸. These may be the natural targets of the group-1 allergens.

Fourth, if we recognize group-1 allergens as a new group of expansins, then a number of closely related GenBank entries become immediate suspects as wall-loosening proteins. We have named this group the β -expansin family, to distinguish it from the original group of expansins, which are now called α -expansins. The β -expansin family comprises the group-1 allergens and related genes expressed in vegetative tissues (Fig. 2), including *CIM1*, a cytokinin-induced

message identified in soybean cell cultures^{39,40}. The sequences of β -expansins are highly represented in the rice and maize expressed sequence tag (EST) databases, where at last count 150 maize EST sequences could be classified as β -expansins encoded by at least 19 different genes (Y. Wu and D.J.C., unpublished data). In comparison, only a single *Arabidopsis thaliana* EST is considered to be a β -expansin. If *Arabidopsis* is representative of other dicotyledons, this comparison implies an exceptional diversity and abundance of β -expansin genes in grasses such as maize and rice. We suspect that the evolutionary radiation of β -expansin genes in grasses is linked to the evolutionary divergence of the grass cell wall⁷.

Finally, because the two families of expansins have similar wall-loosening activity, we infer that the key structural and catalytic residues are to be found in the very limited conservation (~20% identity) between α - and β -expansins. Scrutiny of these regions leads to some new guesses about how these proteins might loosen the cell wall, as discussed below.

A role in fruit softening

During ripening of many fruits, wall polysaccharides are hydrolysed by a suite of digestive enzymes whose genes are specifically expressed during the final stages of fruit development⁴¹. For many years these polysaccharide hydrolases were thought to cause softening of fruit⁴². These ideas were cast into doubt, however, when plants were genetically engineered to alter enzyme activity and polysaccharide breakdown in the fruit, but fruit softening occurred more or less normally.

Following these studies, an α -expansin gene *Le-EXP1* was found to be expressed specifically in the later stages of tomato fruit ripening and its expression was also stimulated by the ripening hormone ethylene⁴⁶. These observations suggested that expansins may function in fruit softening. To test this idea, expression of the *Le-EXP1* gene was reduced by antisense methods; the resulting fruit indeed proved to be firmer throughout ripening⁴⁷. Similarly, overexpression of *Le-EXP1* resulted in softer fruit throughout ripening. These textural effects were relatively small, however, and fruit softening still occurred; it was simply delayed or accelerated. Thus, while *Le-EXP1* expression is involved in tomato fruit softening, it is not the sole determinant of this process.

In tomato plants where expression of the *Le-EXP1* gene was driven with a strong constitutive promoter (CaMV 35S), the fruits were smaller and more rubbery but vegetative growth was normal⁴⁷. The lack of effect on vegetative growth indicates either that the *Le-EXP1* protein does not have a cell-growth function or that the plants were able to compensate for ectopic expression of *Le-EXP1* in vegetative tissues.

In other fruits, such as strawberry and cantaloupe, α -expansin mRNAs are also expressed in late stages of ripening^{46,48} and so this may be a common feature of fruit softening⁴¹. Wall protein extracts from several kinds of ripening fruit possess significant expansin activity⁴⁹, although such activity was not detected in wall proteins from tomato fruit. This is another hint that *Le-EXP1* may lack wall-extension activity. However, it is also possible that much of the extracted protein was inactivated either during extraction or while still within the cell wall. There is precedent for expansin inactivation in growing cucumber seedlings, where expansins are secreted in an active state but become inactivated after the growing cells are displaced from the growth zone. The molecular basis for this inactivation is unclear, but it may relate to thiol oxidation^{11,50}.

More genes, more biological roles

Thanks to the plant genome projects, we now have a nearly complete inventory of the genes in *Arabidopsis* and an extensive, but still incomplete, inventory of expressed genes in rice, maize and other plants. This has provided us with some surprises about expansins. As mentioned above, we have learned that the β -expansin gene family is large and abundantly expressed in grasses (rice and maize), whereas in *Arabidopsis* this gene family is much smaller and not highly

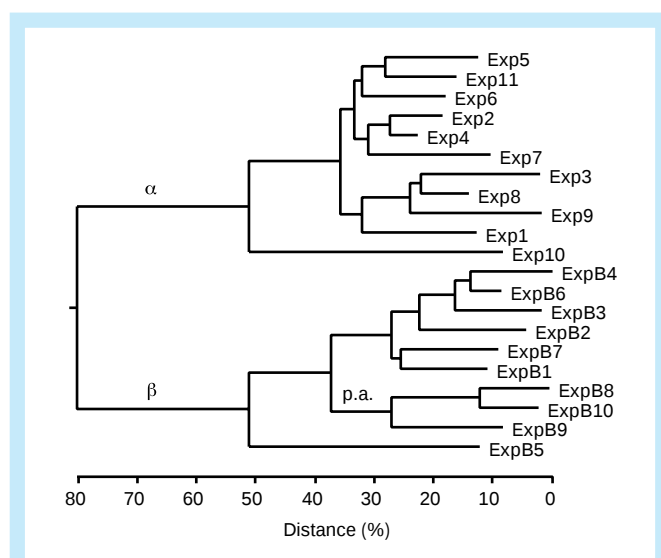


Figure 2 Phylogenetic tree of rice α - and β -expansins. The α - and β -expansins share only about 20% amino-acid identity. Rice appears to have three group-1 pollen allergens (p.a.) that form a distinct subclass of β -expansins. The other β -expansins are expressed in tissues other than pollen and so are referred to as 'vegetative' β -expansins.

expressed. In comparison, *Arabidopsis* has at least 24 α -expansin genes (see the expansin web site at www.bio.psu.edu/expansins/ for an up-to-date listing).

There is evidence that many of these genes have cell-specific expression⁵¹, with each cell type expressing a distinct subset of expansins during growth; for example, trichomes, guard cells, root hairs and xylem each express a distinct set of expansins (ref. 51 and unpublished data). In some tissues, such as the xylem, at least three different expansin genes are expressed, indicating that there may be functional overlap and genetic redundancy. Expansin mRNAs in *Zinnia* xylem are localized to the ends of the cells, where they may function for targeted secretion of expansin to specific walls⁵².

The wealth of GenBank data also lets us evaluate the evolutionary distribution of expansins. So far, they are restricted to land plants (Embryophyta) and are not reported in other kingdoms. In particular, the completed model genomes for bacteria, fungi, insects and nematodes *Escherichia coli*, *Saccharomyces cerevisiae*, *Drosophila melanogaster* and *Caenorhabditis elegans* lack anything that resembles expansin. Because these organisms make a very different kind of cell wall or extracellular matrix compared with that of plants, this finding is consistent with expansin's unique function. However, one might expect to find expansin-like proteins in organisms that invade plant tissues or digest cell walls, such as termites, nematodes that form root cysts, plant parasites and pathogenic fungi. In this regard, it is noteworthy that the digestive track of garden snails *Helix pomatia* contains expansin-like proteins²⁵. These presumably help digestion of plant tissues in the snail's diet. That this is the only documented case of expansin-like proteins in organisms other than plants probably reflects our lack of comprehensive information.

Characteristics of wall loosening by expansin

Expansin's unique physical effects on plant cell walls include rapid induction of wall extension and stimulation of stress relaxation¹¹. Expansins do not progressively weaken the cell wall, nor do they cause a lasting change in wall structure⁵³ (except, of course, that the wall is longer and thinner after it extends). No ligands or cofactors have been identified as necessary for expansin action, although thiol reductants help to maintain stable activity.

Normally, expansin is a very minor component of the cell wall²⁴; for instance, in cucumber seedlings expansin is found at roughly one part protein to 5,000 parts cell wall (on a dry mass basis) and induces wall extension when added in amounts as low as 1:10,000. Binding and activity both saturate at a protein-to-wall ratio of about 1:1,000.

From the architecture of the plant cell wall, one might expect expansin's loosening action to result from hydrolysis of the matrix polymers that hold the cellulose microfibrils in place. However,

attempts to identify hydrolytic activity by expansin against the major wall components have failed^{11,24}. Attempts to mimic expansin's effects with wall hydrolases have also failed²⁵; such enzymes can reduce the mechanical strength of the cell wall, but do not induce extension and stress relaxation, at least not until the wall is degraded to the point of mechanical failure.

A recent study⁵⁴ proposed that group-1 allergens are cysteine proteases, based on proteolytic activity associated with recombinant protein. Spurred by this report, we have examined *Zea m1* and other group-1 allergens for proteolytic activity, but have not detected such activity (L.-C. Li and D.J.C., manuscript in preparation). We also tested a broad range of proteases¹⁰, but failed to detect significant wall-extension activity. Thus, proteolysis does not appear to be a probable mechanism of expansin action.

Model of expansin action

In contrast to these conventional enzymatic theories of wall loosening, we have proposed an alternative mechanism, in which expansins weaken the non-covalent binding between wall polysaccharides, thereby allowing turgor-driven polymer creep⁵⁵ (Fig. 3). In this scheme, expansin would make use of the mechanical strain energy in the wall to catalyse an inchworm-like movement, or reptation, of the wall polymers. Expansin movement may be confined to lateral diffusion along the surface of the cellulose microfibril, as been observed for other polysaccharide-binding proteins⁵⁶. Such contained diffusion would enable expansin to search the microfibril surface, locally loosening its attachment to the matrix, and allowing chain movement and stress relaxation.

This model of expansin action is consistent with the lack of progressive weakening of the cell wall by expansin and also fits the biophysical characteristics of cell-wall extension, the rate of which is dependent on the wall stress in excess of a minimum yield threshold⁵⁷. Other support for this mechanism comes from the finding that expansin weakens pure cellulose paper, whose strength derives from non-covalent binding between cellulose fibres⁵⁵, and artificial composites composed of cellulose alone or cellulose and xyloglucan⁵⁸. Finally, although expansin lacks hydrolytic activity by itself, it does enhance the hydrolysis of crystalline cellulose by cellulases⁵⁹. This synergistic action may indicate that expansin makes glucans on the surface of the microfibril more accessible to enzymatic attack by cellulases.

Sequence analysis of expansins

Although the protein structure for expansin has not yet been solved, sequence analysis hints at a two-domain structure (Fig. 4a). The carboxy terminus is up to 50% identical to a small protein produced by grass pollen, known as a group-2 grass pollen allergen. The biologi-

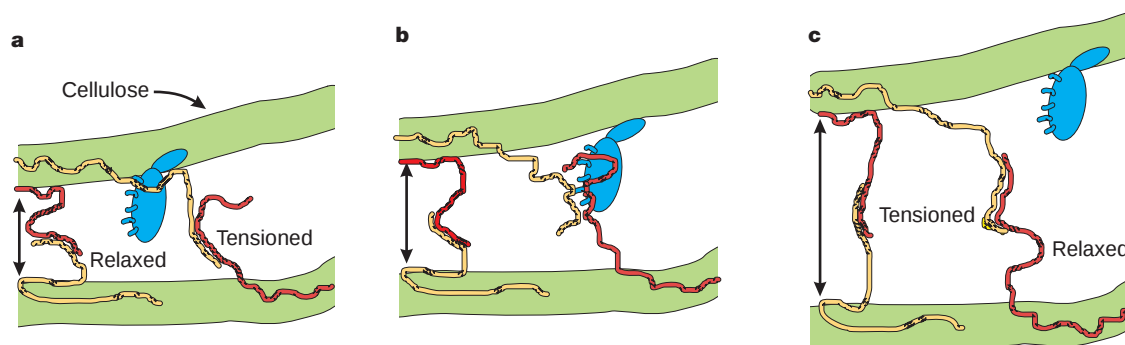


Figure 3 A model of expansin's wall-loosening action. Cellulose microfibrils are connected to each other by glycans (thin yellow and red strands) that can stick to the microfibril surface and to each other. The expansin protein (blue) is hypothesized to disrupt the bonding of the glycans to the microfibril surface (a) or to each other (b). Under the mechanical stress arising from turgor, expansin action results in a displacement of the wall polymers (c) and slippage in the points of polymer adhesion (compare b and c).

cal function of group-2 allergens is unknown, but a potential function in wall loosening is suggested by its homology to expansin and by the fact that it is secreted to the cell wall by grass pollen.

Recently the structure of a group-2 allergen was solved^{60,61} and we have used this structure to model the C terminus of the β -expansin Lol p1 (). On the protein surface are a number of aromatic residues that could function in polysaccharide binding, for example, by ring stacking as found for other carbohydrate-binding proteins⁶². However, the surface does not closely resemble the flat binding surface reported for cellulose-binding domains⁶³, which preferentially bind to the crystalline surface of cellulose.

Expansins also have limited sequence similarity to the catalytic domain of a small family of endoglucanases (family 45)⁶⁴. The sequence identity is only around 30%, but it includes a series of cysteines (Fig. 4c), suggesting a conserved protein fold⁵⁰. In addition, the His-Phe-Asp (HFD) motif found in the enzyme's catalytic site⁶⁵ is also well conserved in most expansins. This conservation led us to re-examine expansins for endoglucanase activity, using more sensitive methods, but the results continue to support a nonhydrolytic mechanism of action.

The resemblance of expansin to family-45 endoglucanases remains enigmatic. Perhaps expansins act by a transglycosylation mechanism, in which the backbone of a polysaccharide is cut and ligated to the end of another similar polymer⁶⁶. Tests for transglycosylation of xyloglucan (the major hemicellulose of most growing cell walls) by α -expansins have given negative results⁶⁷. Moreover, xyloglucan endotransglycosylase lacks wall-extension activity. Also weighing against this idea is that family-45 endoglucanases are inverting enzymes, which do not form the covalent intermediates needed for transglycosylation⁶⁵.

Implications for wall structure

Another enigma concerns the model of the plant cell wall (Box 1). One would reasonably expect such a structure to extend under the action of hydrolases or transglycosylases that act on matrix polysaccharides, but this has not proven to be the case.

Similarly, the immediate induction of wall extension by expansin (starting within seconds and reaching a maximum rate within a couple of minutes after addition) and the minuscule amounts of expansin needed for wall extension suggest that there may be very few 'sticky' spots or entanglements at any moment where expansins act to promote polysaccharide slippage. This inference is supported by NMR studies indicating that expansin does not increase the general mobility of wall polysaccharides⁵³. Thus wall enlargement may be controlled by action at relatively rare sites that link microfibrils together. The nature of these linkage sites remains a mystery.

Finally, the two-domain model of expansin structure suggests that the loosening action of expansin may be a separate function from its binding to the wall. Binding could serve to anchor the protein and to prevent its diffusion to neighbouring cells, where it could interfere with cell-specific control of cell enlargement. Such tight binding, however, would not necessarily prevent protein diffusion along the surface of the microfibril; hence, expansin's action could be localized by its wall-binding properties, which may vary among different kinds of expansin.

Perspectives

This class of protein appears to have evolved specifically in the land plant lineage, with biological functions in cell growth and in other situations where the movement, adhesion and enzymatic accessibility of wall polysaccharides are important. Expansins offer potential applications for bioengineering of cell walls, either to manipulate the growth and texture of plants or to modify the structure and physical properties of cell walls used in commercial products such as wood, textiles and polymers.

Although the plant cell wall is composed of polymers characteristic of plants, analogous structures based on different polymers but similar physical principles are found in fungi, arthropods and other organisms. Perhaps they too have evolved proteins functionally analogous to expansin. □

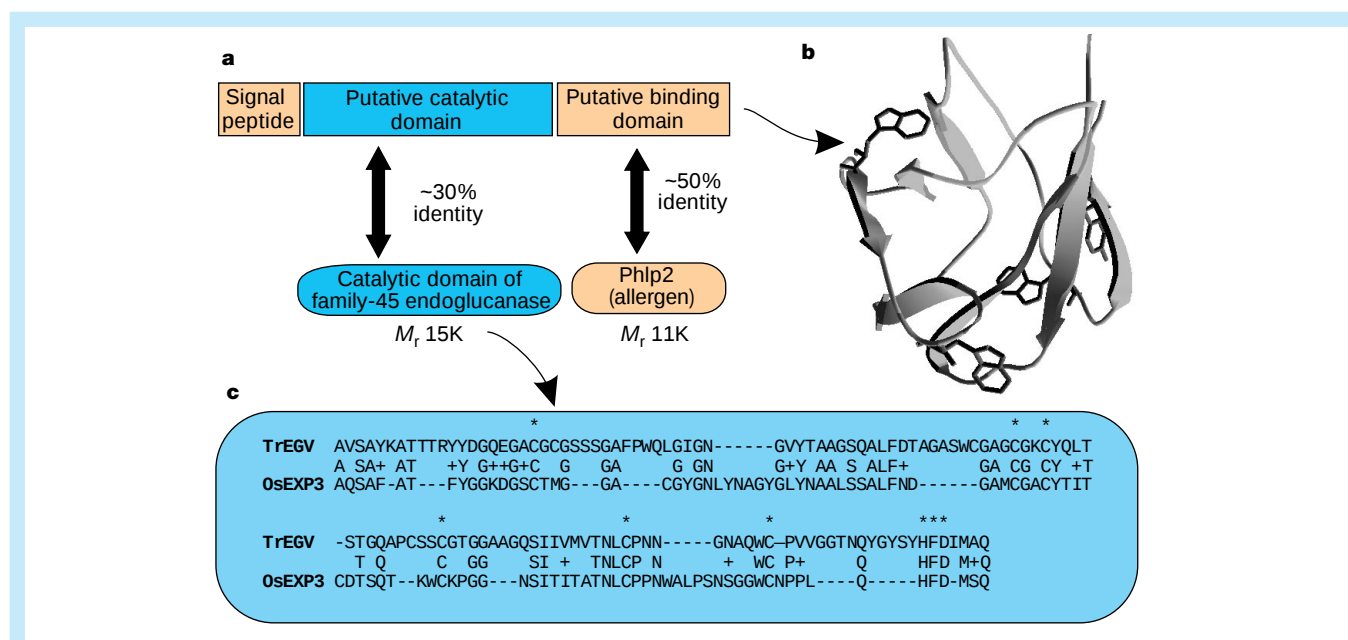


Figure 4 Predicted structure of expansin protein. **a**, Block diagram of putative expansin protein domains, mapping the carboxy terminus to grass pollen group-2 allergens and the amino terminus to the catalytic domain of family-45 endoglucanase. **b**, The structure of the C terminus of Lol p1 (a group-1 pollen allergen and β -expansin), modelled by SWISS-MODEL⁶⁸ using the structure of Phl p2 as a template. Conserved surface aromatic residues (tryptophan, tyrosine) with a possible role in polysaccharide binding are displayed. **c**, Sequence alignment of the catalytic domain of a family-45 endoglucanase (EGV from *Trichoderma reesii*) to a rice α -expansin (Os-EXP3). The endoglucanase also has a linker region and a cellulose-binding domain (not shown) which are not homologous with expansin sequences.

1. Varner, J. E. & Lin, L.-S. Plant cell wall architecture. *Cell* 56, 231–239 (1989).
2. Talbot, L. D. & Ray, P. M. Molecular size and separability features of pea cell wall polysaccharides. Implications for models of primary wall structure. *Plant Physiol.* 92, 357–368 (1992).
3. Roberts, K. The plant extracellular matrix, in a new expansive mood. *Curr. Opin. Cell Biol.* 6, 688–694 (1994).
4. McCann, M. C., Wells, B. & Roberts, K. Complexity in the spatial localization and length distribution of plant cell-wall matrix polysaccharides. *J. Microsc.* 166, 123–136 (1992).
5. McCann, M. C., Wells, B. & Roberts, K. Direct visualization of cross-links in the primary plant cell wall. *J. Cell Sci.* 96, 323–334 (1990).
6. Carpita, N. C. & Gibeau, D. M. Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30 (1993).
7. Cosgrove, D. J. Enzymes and other agents that enhance cell wall extensibility. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 391–417 (1999).
8. Cosgrove, D. J. Water uptake by growing cells: an assessment of the controlling roles of wall relaxation, solute uptake and hydraulic conductance. *Int. J. Plant. Sci.* 154, 10–21 (1993).
9. Rayle, D. L. & Cleland, R. E. The acid growth theory of auxin-induced cell elongation is alive and well. *Plant Physiol.* 99, 1271–1274 (1992).
10. Cosgrove, D. J. Characterization of long-term extension of isolated cell walls from growing cucumber hypocotyls. *Planta* 177, 121–130 (1989).
11. McQueen-Mason, S., Durachko, D. M. & Cosgrove, D. J. Two endogenous proteins that induce cell wall expansion in plants. *Plant Cell* 4, 1425–1433 (1992).
12. Cosgrove, D. J. & Li, Z.-C. Role of expansin in developmental and light control of growth and wall extension in oat coleoptiles. *Plant Physiol.* 103, 1321–1328 (1993).
13. Keller, E. & Cosgrove, D. J. Expansins in growing tomato leaves. *Plant J.* 8, 795–802 (1995).
14. Cho, H. T. & Kende, H. Expansins and internodal growth of deepwater rice. *Plant Physiol.* 113, 1145–1151 (1997).
15. Wu, Y., Sharp, R. E., Durachko, D. M. & Cosgrove, D. J. Growth maintenance of the maize primary root at low water potentials involves increases in cell-wall extension properties, expansin activity, and wall susceptibility to expansins. *Plant Physiol.* 111, 765–772 (1996).
16. Cho, H. T. & Kende, H. Expression of expansin genes is correlated with growth in deepwater rice. *Plant Cell* 9, 1661–1671 (1997).
17. Vriezen, W. H., De Graaf, B., Mariani, C. & Vosenek, L. A. C. J. Submergence induces expansin gene expression in flooding tolerant *Rumex palustris* and not in flooding intolerant *R. acetosa*. *Planta* 210, 956–963 (2000).
18. Hutchison, K. W., Singer, P. B., Diaz-Sala, C. & Greenwood, M. S. Expansins are conserved in conifers and expressed in response to exogenous auxin. *Plant Physiol.* 120, 827–832 (1999).
19. Fleming, A. J., Caderas, D., Wehrli, E., McQueen-Mason, S. & Kuhlmeier, C. Analysis of expansin-induced morphogenesis on the apical meristem of tomato. *Planta* 208, 166–174 (1999).
20. Moore, R. C., Flecker, D. & Cosgrove, D. J. Expansin action on cells with tip growth and diffuse growth. *J. Cell. Biochem. (Suppl.)* 21A, 457; abstract J5–312 (1995).
21. Link, B. M. & Cosgrove, D. J. Acid-growth response and α -expansins in suspension cultures of bright yellow 2 tobacco. *Plant Physiol.* 118, 907–916 (1998).
22. Fleming, A. J., McQueen-Mason, S., Mandel, T. & Kuhlmeier, C. Induction of leaf primordia by the cell wall protein expansin. *Science* 276, 1415–1418 (1997).
23. Green, P. B. Expansin and morphology: a role for biophysics. *Trends Plant Sci.* 2, 365–366 (1997).
24. McQueen-Mason, S. & Cosgrove, D. J. Expansin mode of action on cell walls: Analysis of wall hydrolysis, stress relaxation, and binding. *Plant Physiol.* 107, 87–100 (1995).
25. Cosgrove, D. J. & Durachko, D. M. Autolysis and extension of isolated walls from growing cucumber hypocotyls. *J. Exp. Bot.* 45, 1711–1719 (1994).
26. Reinhardt, D., Wittwer, F., Mandel, T. & Kuhlmeier, C. Localized upregulation of a new expansin gene predicts the site of leaf formation in the tomato meristem. *Plant Cell* 10, 1427–1437 (1998).
27. Kende, H., Van der Knaap, E. & Cho, H. T. Deepwater rice: A model plant to study stem elongation. *Plant Physiol.* 118, 1105–1110 (1998).
28. Catala, C., Rose, J. K. & Bennett, A. B. Auxin-regulated genes encoding cell wall-modifying proteins are expressed during early tomato fruit growth. *Plant Physiol.* 122, 527–534 (2000).
29. Brummell, D. A., Harpster, M. H. & Dunsmuir, P. Differential expression of expansin gene family members during growth and ripening of tomato fruit. *Plant Mol. Biol.* 39, 161–169 (1999).
30. Orford, S. J. & Timmis, J. N. Specific expression of an expansin gene during elongation of cotton fibres. *Biochim. Biophys. Acta Gene Struct. Expression* 1398, 342–346 (1998).
31. Shimizu, Y. et al. Changes in levels of mRNAs for cell wall-related enzymes in growing cotton fiber cells. *Plant Cell Physiol.* 38, 375–378 (1997).
32. Cho, H. T. & Cosgrove, D. J. Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA* (in the press).
33. Shcherban, T. Y. et al. Molecular cloning and sequence analysis of expansins—a highly conserved, multigene family of proteins that mediate cell wall extension in plants. *Proc. Natl Acad. Sci. USA* 92, 9245–9249 (1995).
34. Smith, P. M., Knox, R. B. & Singh, M. B. in *Pollen Biotechnology: Gene Expression and Allergen Characterization* (eds Mohapatra, S. S. & Knox, R. B.) 125–143 (Chapman & Hall, New York, 1996).
35. Knox, B. & Suphioglu, C. Environmental and molecular biology of pollen allergens. *Trends Plant Sci.* 1, 156–164 (1996).
36. Cosgrove, D. J., Bedinger, P. & Durachko, D. M. Group I allergens of grass pollen as cell wall-loosening agents. *Proc. Natl Acad. Sci. USA* 94, 6559–6564 (1997).
37. Heslop-Harrison, Y., Reger, B. J. & Heslop-Harrison, J. The pollen-stigma interaction in the grasses. 5. Tissue organization and cytochemistry of the stigma (“silk”) of *Zea mays* L. *Acta Bot. Neerl.* 33, 81–99 (1984).
38. Carpita, N. C. Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 445–476 (1996).
39. Crowell, D. N. Cytokinin regulation of a soybean pollen allergen gene. *Plant Mol. Biol.* 25, 829–835 (1994).
40. Downes, B. P. & Crowell, D. N. Cytokinin regulates the expression of a soybean β -expansin gene by a post-transcriptional mechanism. *Plant Mol. Biol.* 37, 437–444 (1998).
41. Rose, J. K. C. & Bennett, A. B. Cooperative disassembly of the cellulose-xyloglucan network of plant cell walls: parallels between cell expansion and fruit ripening. *Trends Plant Sci.* 4, 176–183 (1999).
42. Fischer, R. L. & Bennett, A. B. Role of cell wall hydrolases in fruit ripening. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 42, 675–703 (1991).
43. Giovannoni, J. J., DellaPenna, D., Bennett, A. B. & Fischer, R. L. Expression of a chimeric polygalacturonase gene in transgenic rin (ripening inhibitor) tomato fruit results in polyuronide degradation but not fruit softening. *Plant Cell* 1, 53–63 (1989).
44. Tieman, D. M., Harriman, R. W., Ramamohan, G. & Handa, A. K. An antisense pectin methylesterase gene alters pectin chemistry and soluble solids in tomato fruit. *Plant Cell* 4, 667–679 (1992).
45. Brummell, D. A., Hall, B. D. & Bennett, A. B. Antisense suppression of tomato endo-1,4- β -glucanase Cel2 mRNA accumulation increases the force required to break fruit abscission zones but does not affect fruit softening. *Plant Mol. Biol.* 40, 615–622 (1999).
46. Rose, J. K. C., Lee, H. H. & Bennett, A. B. Expression of a divergent expansin gene is fruit-specific and ripening-regulated. *Proc. Natl Acad. Sci. USA* 94, 5955–5960 (1997).
47. Brummell, D. A. et al. Modification of expansin protein abundance in tomato fruit alters softening and cell wall polymer metabolism during ripening. *Plant Cell* 11, 2203–2216 (1999).
48. Civello, P. M., Powell, A. L. T., Sabehat, A. B. & Bennett, A. B. An expansin gene expressed in ripening strawberry fruit. *Plant Physiol.* 121, 1273–1279 (1999).
49. Rose, J. K. C., Cosgrove, D. J., Albersheim, P., Darvill, A. G. & Bennett, A. B. Detection of expansin proteins and activity during tomato fruit ontogeny. *Plant Physiol.* (in the press).
50. Cosgrove, D. J. Relaxation in a high-stress environment: The molecular bases of extensible cell walls and cell enlargement. *Plant Cell* 9, 1031–1041 (1997).
51. Durachko, D. M. & Cosgrove, D. J. Expression patterns for selective expansin genes in *Arabidopsis*. *Annu. Meeting Am. Soc. Plant Physiol. Abstr.* 56 (1999).
52. Im, K. H., Cosgrove, D. J. & Jones, A. M. Subcellular localization of expansin mRNA in xylem cells. *Plant Physiol.* 123, 463–470 (2000).
53. Fenwick, K. M., Apperley, D. C., Cosgrove, D. J. & Jarvis, M. C. Polymer mobility in cell walls of cucumber hypocotyls. *Phytochem.* 51, 17–22 (1999).
54. Grobe, K., Becker, W. M. & Petersen, A. Grass group I allergens (β -expansins) are novel, papain-related proteinases. *Eur. J. Biochem.* 263, 33–40 (1999).
55. McQueen-Mason, S. & Cosgrove, D. J. Disruption of hydrogen bonding between wall polymers by proteins that induce plant wall extension. *Proc. Natl Acad. Sci. USA* 91, 6574–6578 (1994).
56. Jervis, E. J., Haynes, C. A. & Kilburn, D. G. Surface diffusion of cellulases and their isolated binding domains on cellulose. *J. Biol. Chem.* 272, 24016–24023 (1997).
57. Cosgrove, D. J. Biophysical control of plant cell growth. *Annu. Rev. Plant Physiol.* 37, 377–405 (1986).
58. Whitney, S. E. C., Gidley, M. J. & McQueen-Mason, S. Probing expansin action using cellulose/hemicellulose composites. *Plant J.* 22, 327–334 (2000).
59. Cosgrove, D. J., Durachko, D. M. & Li, L.-C. Expansins may have cryptic endoglucanase activity and can synergize the breakdown of cellulose by fungal cellulases. *Annu. Meeting Am. Soc. Plant Physiol. Abstr.* 171 (1998).
60. De Marino, S. et al. An immunoglobulin-like fold in a major plant allergen: the solution structure of Phl p 2 from timothy grass pollen. *Struct. Fold Des.* 7, 943–952 (1999).
61. Fedorov, A. A., Ball, T., Valenta, R. & Almo, S. C. X-ray crystal structures of birch pollen profilin and Phl p 2. *Int. Arch. Allergy Immunol.* 113, 109–113 (1997).
62. Toone, E. J. Structure and energetics of protein-carbohydrate complexes. *Curr. Opin. Struct. Biol.* 4, 719–728 (1994).
63. Mattinen, M. L., Linder, M., Drakenberg, T. & Annala, A. Solution structure of the cellulose-binding domain of endoglucanase I from *Trichoderma reesei* and its interaction with cello-oligosaccharides. *Eur. J. Biochem.* 256, 279–286 (1998).
64. Henrissat, B., Teeri, T. T. & Warren, R. A. J. A scheme for designating enzymes that hydrolyse the polysaccharides in the cell walls of plants. *FEBS Lett.* 425, 352–354 (1998).
65. Davies, G. J. et al. Structure and function of endoglucanase V. *Nature* 365, 362–364 (1993).
66. Fry, S. C. Polysaccharide-modifying enzymes in the plant cell wall. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 46, 497–520 (1995).
67. McQueen-Mason, S., Fry, S. C., Durachko, D. M. & Cosgrove, D. J. The relationship between xyloglucan endotransglycosylase and in vitro cell wall extension in cucumber hypocotyls. *Planta* 190, 327–331 (1993).
68. Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis* 18, 2714–2723 (1997).

Supported by the US National Science Foundation, the Department of Energy and the Department of Agriculture.

Correspondence should be addressed to D.J.C. (e-mail: dCosgrove@psu.edu).

Structure of the 30S ribosomal subunit

Brian T. Wimberly^{*†}, Ditlev E. Brodersen^{*†}, William M. Clemons Jr^{*†‡}, Robert J. Morgan-Warren^{*†}, Andrew P. Carter^{*†}, Clemens Vornrhein[§], Thomas Hartsch^{||} & V. Ramakrishnan[†]

[†]MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

[‡]Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, Utah 84132, USA

[§]Global Phasing Ltd., Sheraton House, Castle Park, Cambridge CB3 0AX, UK

^{||}Göttingen Genomics Laboratory, Institut für Mikrobiologie und Genetik, Georg-August-Universität Göttingen, Grisebachstr. 8, D-37077 Göttingen, Germany

* These authors contributed equally to this work

Genetic information encoded in messenger RNA is translated into protein by the ribosome, which is a large nucleoprotein complex comprising two subunits, denoted 30S and 50S in bacteria. Here we report the crystal structure of the 30S subunit from *Thermus thermophilus*, refined to 3 Å resolution. The final atomic model rationalizes over four decades of biochemical data on the ribosome, and provides a wealth of information about RNA and protein structure, protein–RNA interactions and ribosome assembly. It is also a structural basis for analysis of the functions of the 30S subunit, such as decoding, and for understanding the action of antibiotics. The structure will facilitate the interpretation in molecular terms of lower resolution structural data on several functional states of the ribosome from electron microscopy and crystallography.

Protein synthesis is a complex, multistep process that requires, in addition to the ribosome, several extrinsic GTP-hydrolysing protein factors during each of the main stages of initiation, elongation and termination. The 30S ribosomal subunit has a crucial role in decoding mRNA by monitoring base pairing between the codon on mRNA and the anticodon on transfer RNA; the 50S subunit catalyses peptide-bond formation. Despite several decades of work, the molecular details of the process are poorly understood, and the elucidation of the mechanism of translation is one of the fundamental problems in molecular biology today. A recent collection of articles summarizes the field¹.

An important contribution was made by Yonath and co-workers², who showed that structures as large as the 50S ribosomal subunit would form crystals that diffract beyond 3 Å resolution. Originally it was not clear that phase information from such a large asymmetric unit could be obtained to high resolution, but the development of bright, tunable synchrotron radiation sources, large and accurate area detectors, vastly improved crystallographic computing, and the advent of cryocrystallography have all contributed to making structural studies of the ribosome more tractable. In our work, the use of anomalous scattering from the LIII edges of lanthanides and osmium has also played a critical role in obtaining phases³.

The 30S ribosomal subunit (hereafter referred to as 30S) from *Thermus thermophilus* was originally crystallized by the Puschino group in 2-methyl-2,4-pentenediol (MPD)⁴, and by Yonath and co-workers⁵ in a mixture of ethyl-butanol and ethanol. The MPD crystal form originally diffracted to about 9–12 Å resolution^{6,7}. The diffraction limit of these crystals did not improve beyond 7 Å resolution for almost a decade, but more recently both Yonath *et al.*^{8,9} and we³ obtained crystals of the MPD form that exhibit significantly improved diffraction. However, unlike the crystals obtained by the Yonath group⁹, our crystals do not require soaking in tungsten clusters or heat treatment to obtain high-resolution diffraction.

Last year, we described the structure of the 30S at 5.5 Å resolution³. We placed all seven proteins whose structures were known at the time, inferred the structure of protein S20 to be a three-helix bundle, traced the fold of an entire domain of 16S RNA, and identified a long RNA helix at the interface that contains the decoding site of the 30S. Proteins S5 and S7 were also placed in electron density maps of the 30S obtained by Yonath *et al.*⁹. We have now solved and refined the structure of the 30S at 3 Å resolution. The structure contains all of the ordered regions of 16S RNA and 20 associated proteins, constituting over 99% of the total 16S RNA and

95% of the ribosomal proteins, with the missing parts being exclusively at the termini of RNA or polypeptide chains. Here we describe the overall structural organization of the 30S subunit, and in an accompanying paper¹⁰ we describe functional insights gleaned from the structure and the interaction of antibiotics bound to the 30S. A more detailed analysis of the structure will be presented elsewhere.

Results

Crystallographic statistics are presented in Table 1. Experimentally phased maps clearly showed main-chain density for RNA and protein, individual bases (of sufficient quality to distinguish purines from pyrimidines) and large well-ordered side chains of proteins (Fig. 1). The structure was initially built from these experimental maps and was rebuilt after a round of refinement. The current model consists of nucleotides 5–1,511 of *Thermus thermophilus* 16S RNA (corresponding to 5–1,534 of *Escherichia coli* 16S RNA)¹¹ and all of the ordered regions of the associated 20 proteins. These proteins correspond to *E. coli* proteins S2–S20 and a small 26-residue peptide, Thx (ref. 12). *Thermus* does not contain S21, and in our work S1 was removed from the 30S before crystallization. The model has been refined against 3.05 Å native data, resulting in an R/R_{free} of 0.208/0.252 with good geometry. For the proteins, 95.7% of the residues were in the core or allowed regions of the Ramachandran plot, 2.4% in the generously allowed region and 1.9% in the disallowed region.

Overview of the 30S

The secondary structure diagram for 16S RNA is shown in Fig. 2a, along with the definitions for the standard helix numbering H1–H45 (ref. 13). The sequence numbering for *E. coli* is used throughout; the main difference is that *Thermus* has a shorter H6 and H10, and insertions in H9 and H33a. Insertions in *Thermus* relative to *E. coli* are indicated in the coordinates by a letter code following the practice for transfer RNA.

The overall shape of the 30S is very similar to the model derived from negatively stained electron microscopy samples and to more recent cryo-electron microscopy reconstructions¹⁴, and appears to be closer to the 50S-bound form than the different free 30S form¹⁵. The shape is largely determined by the RNA component; none of the gross morphological features is all protein. In the canonical ‘front’ view from the 50S, the tertiary fold of 16S RNA (Fig. 2b) shows the head with a beak pointing leftwards, the body with the shoulder at top left and the spur at lower left, and the platform at top right.

Individual secondary structure domains (Fig. 2a) make up each of these morphological features (Fig. 2b), consistent with proposals made in previous modelling studies^{13,16,17}. The 5' domain makes up the bulk of the body; the central domain most of the platform, and the 3' major domain constitutes the bulk of the head. The 3' minor domain is the only significant exception to this rule, as it is part of the body at the subunit interface. The four domains of the 16S RNA secondary structure radiate from a central point in the neck region of the subunit, and are especially tightly associated in this area, which is functionally the most important region of the 30S ribosomal subunit.

The distribution of proteins and RNA in the 30S is asymmetric, as was predicted from neutron scattering¹⁸. The proteins are concentrated in the top, sides and back of the 30S (Fig. 2c, d). None of the proteins binds entirely inside an RNA domain, although S20 binds between two domains (the 3' minor domain and 5' domain). The 50S interface is largely free of protein, with the exception of S12 which lies near the decoding site at the top of the long H44 that runs down the interface. Other proteins lie at the periphery of the subunit interface, allowing them to make contact with the 50S subunit. A movie of the structure is available in the Supplementary Information.

The structure of the RNA. The secondary structure of 16S RNA contains over 50 regular double helices connected by irregular single-stranded loops (Fig. 2a). In the crystal structure, many of these formally single-stranded loop regions are in fact only slightly irregular double-stranded extensions of neighbouring regular

helices. Thus, most of 16S RNA may be described as helical or approximately helical, and it is useful to consider the RNA structure as a three-dimensional arrangement of helical elements. Interactions between helical elements include vertical co-axial stacking of helices neighbouring in sequence, and horizontal packing of helices, usually between their minor grooves. Co-axial stacking of helices is very common: the helices of 16S RNA are organized into 13 groups of co-axially stacked helices and 23 unstacked helices, for a total of 36 helical elements. The packing of these helical elements largely determines the overall fold of each of the four domains of 16S RNA. Short single-stranded RNA segments make idiosyncratic long-range interactions to stabilize the packing of helical elements. Proteins also help stabilize the RNA tertiary structure by binding to two or more RNA helical elements, as described below.

Helix packing interactions. There are three types of helix–helix packing in the structure, all of which use the wide and shallow minor groove as an interaction surface (Fig. 3). The most common packing mode is docking of the minor grooves of two helices, as shown for the interaction of H6 with H8 (Fig. 3a). Usually one or both helices are distorted from canonical A-form helical geometry in order to create a larger and more complementary interaction surface. Such helical distortions are caused by the base-pair geometries of noncanonical base pairs and sometimes by a bulged-out base. Noncanonical pairs involving adenines, especially sheared G·A, A·A, and reverse-Hoogsteen U·A and C·A pairs, are particularly common and are often adjacent to create a cross-strand adenine stack motif. This motif widens the minor groove to make it more

Table 1 Crystallographic statistics

Data collection

Data set*	Beamline	Resolution limit (Å)	No. of observations	No. of unique reflections	Completeness (%)		$\langle I \rangle / \langle \sigma \rangle$		R_{sym} (%)	
					Overall	Outer shell	Overall	Outer shell	Overall	Outer shell
Native	ID14-4+SBC 19ID	3.05	840101	254607	94.0	81.7	12.0	1.8	10.8	49.0
Os1	SBC 19IF	3.35	1355593	199999	98.4	90.3	16.4	2.4	13.9	59.2
Os2	ID14-4	4.00	362452	118494	97.4	97.1	8.5	2.3	10.9	38.9
Os3a	ID14-4	4.30	259175	93894	96.4	99.0	6.4	1.7	18.7	53.2
Os3b	SBC 19ID	3.50	714774	175049	97.3	94.7	10.4	2.5	16.7	66.5
Os4	X25	4.50	301216	79267	93.9	65.9	11.6	3.3	9.8	22.9
Lu1a	SBC 19ID	3.35	981691	180035	97.5	83.5	16.9	3.4	10.9	37.3
Lu1b	SBC 19ID	3.20	929154	230296	98.1	97.6	10.0	2.1	14.7	60.5

Phasing

Derivative*	SOLVE†			SHARP‡		
	No. of sites	$R_{\text{cullis}}§$	Phasing power	No. of sites	$R_{\text{cullis}}§$	Phasing power
Os1	56	0.69	0.84	93	0.55	1.15
Os2	6	0.87	0.20	—	—	—
Os3a	49	0.70	0.61	—	—	—
Os3b	—	—	—	23	0.77	0.69
Os4	4	0.94	0.10	—	—	—
Lu1a	18	0.78	0.33	—	—	—
Lu1b	14	0.77	0.39	19	0.83	0.60

Refinement

Resolution range	99–3.05 Å
Reflections excluded for cross-validation	5%
Number of non-hydrogen atoms	
Proteins	19,198
RNA	32,470
Metals	180
R factor (conventional)	0.208
R factor (free)	0.252
Cross-validated coordinate error	0.45 Å
Deviations from ideality	
r.m.s. deviations in bond lengths	0.006 Å
r.m.s. deviations in bond angles	1.24°

* Os1, Osmium hexammine chloride; Os2, Penta-ammine (dinitrogen) osmium(II) chloride; Os3a, b, Penta-ammine (trifluorosulphonato) osmium(III) trifluoromethanesulphonate; Os4, Osmium bipyridine; Lu1a, b, lutetium chloride.

† SOLVE (to 3.35 Å): mean figure of merit: 0.48 (overall); 0.24 (outer shell).

‡ SHARP (to 3.2 Å): mean figure of merit: 0.37 (overall); 0.03 (outer shell).

§ R_{cullis} is for centric reflections.

|| Phasing power is for anomalous differences.

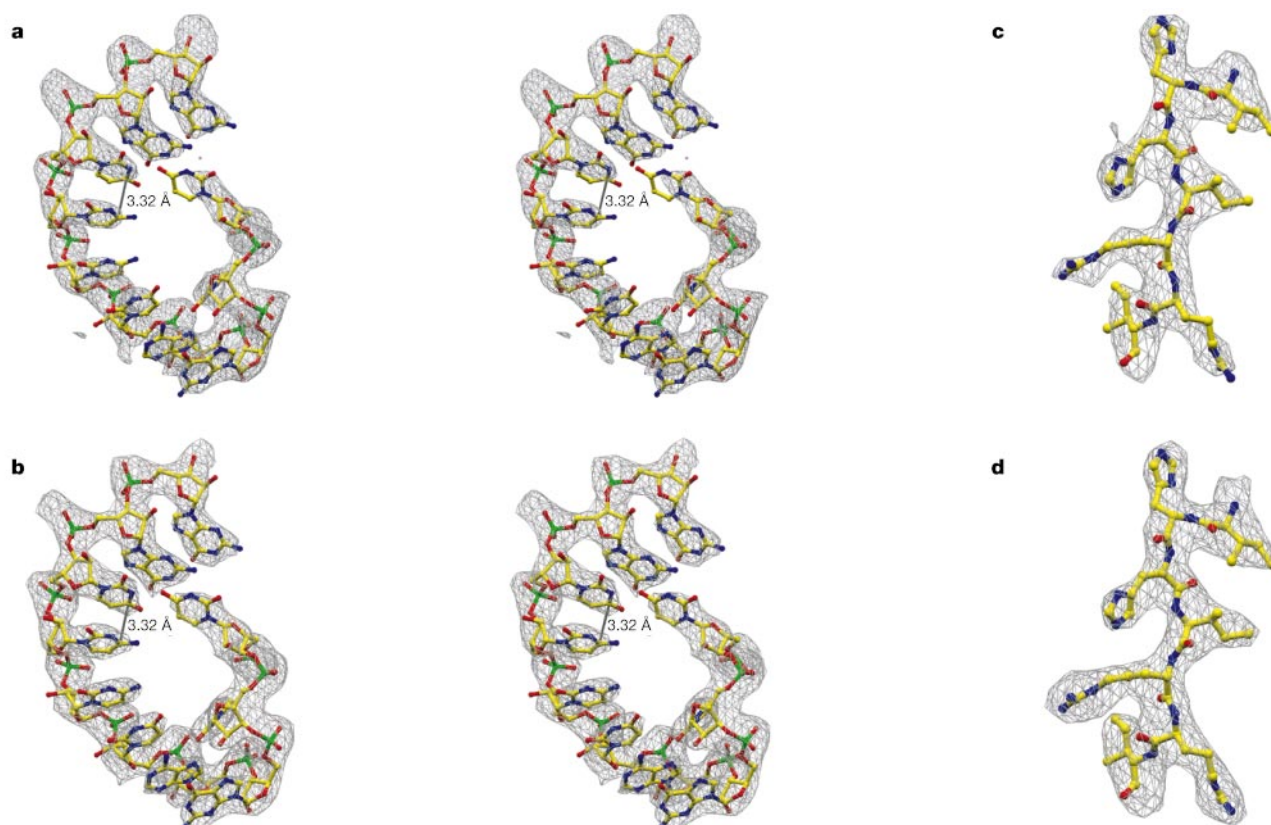


Figure 1 Electron density maps of the 30S. The experimental maps shown here were obtained by phasing with SHARP followed by density modification using SOLOMON and DM (see text); refined maps represent σ_A -weighted ($2mF_o - DF_c$) maps using the final

model. **a, b**, Experimental and refined maps respectively, of a loop of RNA.

c, d, Experimental and refined maps of a strand of protein showing amino-acid side chains. These and other figures were made using RIBBONS³⁷.

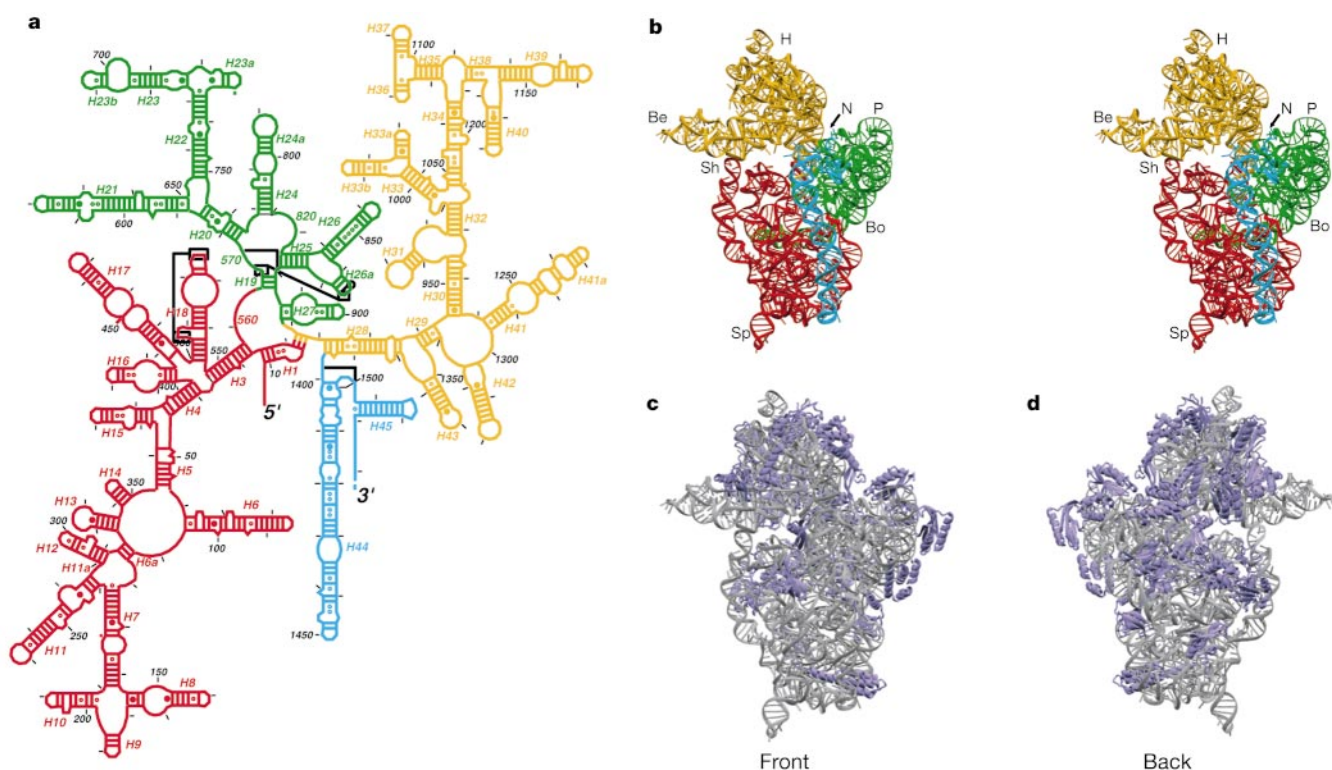


Figure 2 Overview of the 30S structure. **a**, Secondary structure diagram of 16S RNA (modified with permission from <http://www.rna.icmb.utexas.edu/CSI/2STR/Schematics/e.coli16s.27.5.5.schem.ps>; see also ref. 21), showing the definition of the various helical elements used throughout the text. The numbering and diagram correspond to the *E. coli* sequence. Red, 5' domain; green, central domain; orange, 3' major domain; cyan, 3' minor domain. **b**, Stereo view of the tertiary structure of 16S RNA from our refined model, showing the 50S or 'front' view, with the same colouring for the domains. H, head; Be, beak; N, neck; P, platform; Sh, shoulder; Sp, spur; Bo, body. **c, d**, Front (50S) and back sides of the 30S. Grey, RNA; blue, proteins.

minor domain. **b**, Stereo view of the tertiary structure of 16S RNA from our refined model, showing the 50S or 'front' view, with the same colouring for the domains. H, head; Be, beak; N, neck; P, platform; Sh, shoulder; Sp, spur; Bo, body. **c, d**, Front (50S) and back sides of the 30S. Grey, RNA; blue, proteins.

accessible, and it also pushes the adenine bases into the minor groove to facilitate hydrogen bonding to sugar and base functional groups from the packing partner helix. Helix–helix docking results in an extensive and intimate interaction stabilized by a dense network of hydrogen bonds from adenine base nitrogens and 2' OH moieties to the 2' OH, guanine NH₂ and pyrimidine O2 atoms from the partner helix (Fig. 3a, with adenines in red). Some of the distorted adenine-rich structures used to mediate this mode of helix–helix packing are recurrent structural motifs (for example, the common S-turn motif¹⁹). This minor-groove packing mode is not limited to helix–helix interactions: a similar mode is often seen between single-stranded adenines and a regular double helix, particularly the docking of the last three nucleotides of GNRA hairpin loops against the minor groove of a helix. Occasionally a single unpaired adenine base packs against the minor groove of a regular helix.

A second and less common form of helix packing involves the insertion of a ridge of phosphates into the minor groove of another helix. This packing mode is stabilized by hydrogen bonds between the ridge of phosphate oxygens and a layer of 2' OH and guanine base NH₂ groups, as is shown for the interaction between the phosphate backbone of H7 with the minor groove of H21 (Fig. 3b). These guanine NH₂ groups (red in Fig. 3b) are often made more accessible by the geometry of G·U wobble pairs, which places this moiety further into the minor groove compared with

Watson–Crick base pairs. This phosphate-ridge packing mode creates fewer hydrogen bonds and buries less surface area than the minor-groove mode described above, and it may be less stable.

The third and least common mode of helix packing uses an unpaired purine base to mediate the perpendicular packing of one helix against the minor groove of another helix. Three examples of this mode are present in 16S RNA: H27 against the minor groove of a helical stack made up of H1 and H28 (Fig. 3c); H34 against the H35–H34–H38 stack; and H44 against H28. Like the phosphate-ridge mode, this end-on packing may be less stable than the minor-groove packing. Significantly, all of these end-on packing interactions involve functionally important helices, and in one case (H27 against H1/H28) it appears likely that the packing interaction may change as a result of a conformational switch in H27 (ref. 20).

Overview of the domains of 16S RNA

Stacking and packing of the helical elements of 16S RNA generates three compact domains (5' domain, central domain and 3' major domain) and one extended domain (3' minor domain). Packing interactions between the domains create the functionally important 50S and transfer RNA/messenger RNA binding sites. Here we provide an overview of the structure of each domain; details will be published elsewhere.

The 5' domain. The 5' domain is the RNA component of the body. It contains 19 double helices packed as a wedge-shaped mass of RNA

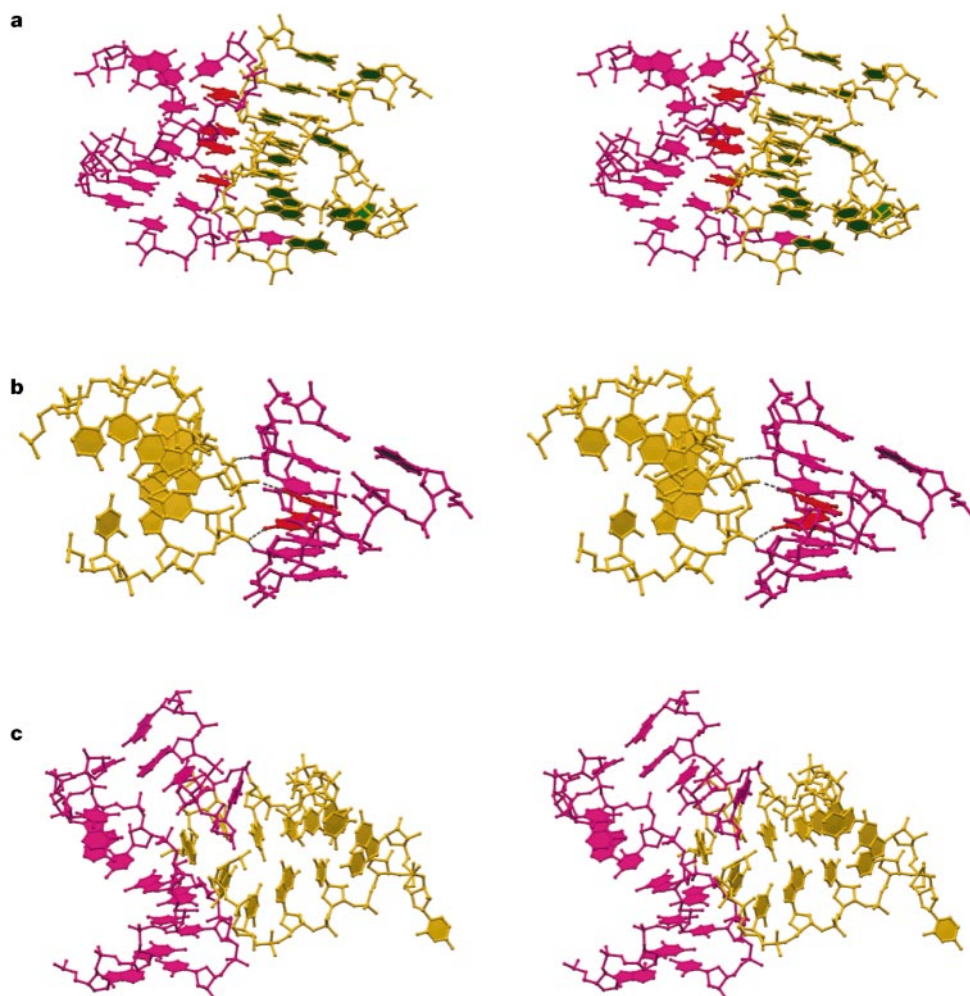


Figure 3 Different modes of interhelical packing in 16S RNA. **a**, The common minor-groove to minor-groove packing mode is often stabilized by a layer of adenosines (red), which mediate most of the hydrogen bonds between two helices (magenta and yellow). **b**, The phosphate ridge to minor groove mode. Usually this mode is stabilized by hydrogen

bonds between guanine N2 groups (red) and phosphate oxygens. **c**, The rare end-on mode of packing uses an unpaired purine (leftmost yellow base) to mediate packing of two helices at right angles to each other.

that tapers to a single layer of double helices near the top (Fig. 4). Like the other domains, it is rather longer along the subunit interface than in the perpendicular direction. The 5' domain can be divided into three subdomains, roughly corresponding to the upper, lower and middle thirds of the secondary structure of the

domain (Fig. 4b–d). These subdomains make up the top and left-hand, the middle and the lower right-hand sides of the body, respectively, in the view from 50S. The spur at the bottom of the 30S is formed by H6, which is known to vary in length across species²¹.

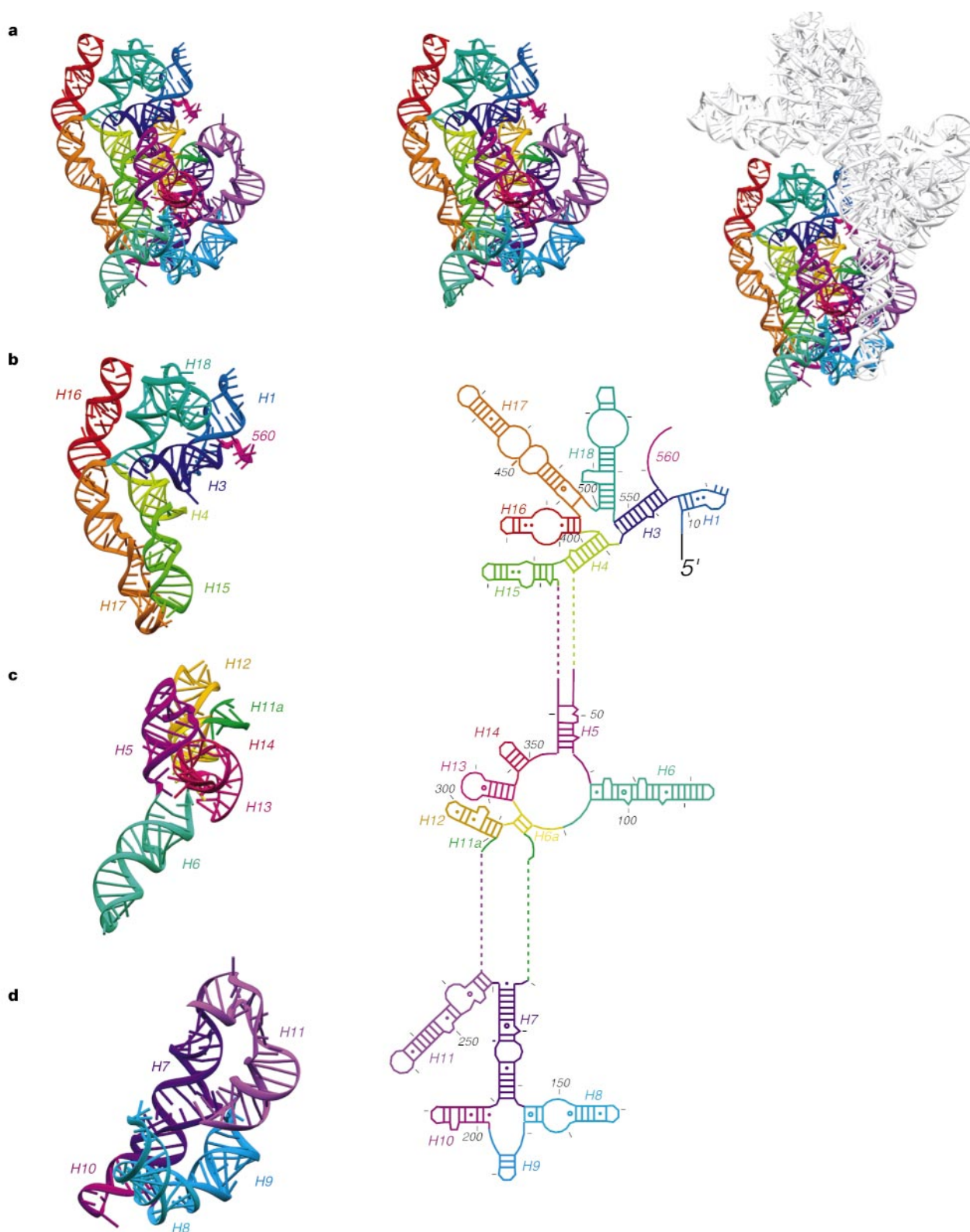


Figure 4 Structure of the 5' domain of 16S RNA. **a**, Stereo view of the entire 5' domain, with an inset on the right showing its location in the 30S subunit. The upper (**b**), middle (**c**) and lower (**d**) subdomains are shown separately next to corresponding parts of the

secondary structure diagrams. The colours in the secondary structure diagrams match those in the structure in this and Figs 5 and 6.

A particularly striking feature is the H16–H17 co-axial stack, which is almost 120 Å long and forms the left-hand border of the body (Fig. 4b), with H16 reaching out to the head. There are also two examples of sharp bends in helices. H18 is sharply bent to accommodate the functionally important 530 pseudoknot (Fig. 4b), which packs against the central pseudoknot at the H18–H1 interface. H11 contains two sharp bends that allow its conserved terminal hairpin loop to pack against H7 (Fig. 4d). Both bends are stabilized by short-range minor-groove to minor-groove packing contacts. Finally, there is an unusual packing interaction between the highly conserved UACG and GAAA tetraloops at the ends of H8 and H14 near the subunit interface.

The central domain. The central domain is the RNA component of the platform. Its fold based on our previous 5.5 Å model³, and the high-resolution structures of parts of it^{22,23} are in excellent agreement with our current structure. It contains nine helical elements folded into a W-shape in the 50S view (Fig. 5). Two long single-stranded segments of RNA, the 570 and 820 loops, are also important structural elements. The central domain is dominated by the long stack of H21–H22–H23, which forms the outer arms of the W. At one end, H21 wraps around the back of the 5' domain; at the other, H22, H23 and the roughly parallel H24 form the bulk of the platform. The tip of the platform consists of H23B and H24A, whose conserved and functionally important hairpin loops (the 690 and 790 loops) are tightly packed. This arrangement requires sharp bends between H23 and H23B, and between H24 and H24A. The H23–H23B bend is stabilized by short-range minor groove/minor

groove packing interactions. The H24–H24A bend is more unusual in that the bend is towards the major groove, which places a ridge of H24A phosphates in the major groove of H24. This major-groove bend is stabilized partly by short-range base–base and base–backbone interactions in the major groove of the bend, and partly by long-range interactions between the bent H24/H24A minor groove and the minor groove of H23.

The 3' major domain. The 3' major domain is the RNA component of the head. The left-hand side of the head tapers to a beak made of RNA on the 50S side and protein on the solvent side (Fig. 6a). The 3' major domain consists of 15 helical elements, which can be roughly divided into three subdomains (Figs 6b–d). The upper subdomain is an extended structure in the part of the head farthest from the 50S subunit, and makes relatively few packing contacts with other RNA. The lower and middle subdomains are more globular and are more intimately packed together, and make up the front-right and front-left portions of the head, respectively. The middle subdomain includes the RNA portion of the beak.

In contrast to the extensive stacking of neighbouring helices seen in the 5' domain and the central domain, most of the helices in the 3' major domain do not stack on a neighbouring helix. An exception is the H35–H36–H38–H39 stack that dominates the upper subdomain (Fig. 6b) and stretches from the top to the bottom of the head. Significantly, the functionally important helices H31 and H34 are quite irregular and make only rather weak packing interactions with other RNA helices (Fig. 6a, c, d).

The 3' minor domain. The 3' minor domain consists of just two

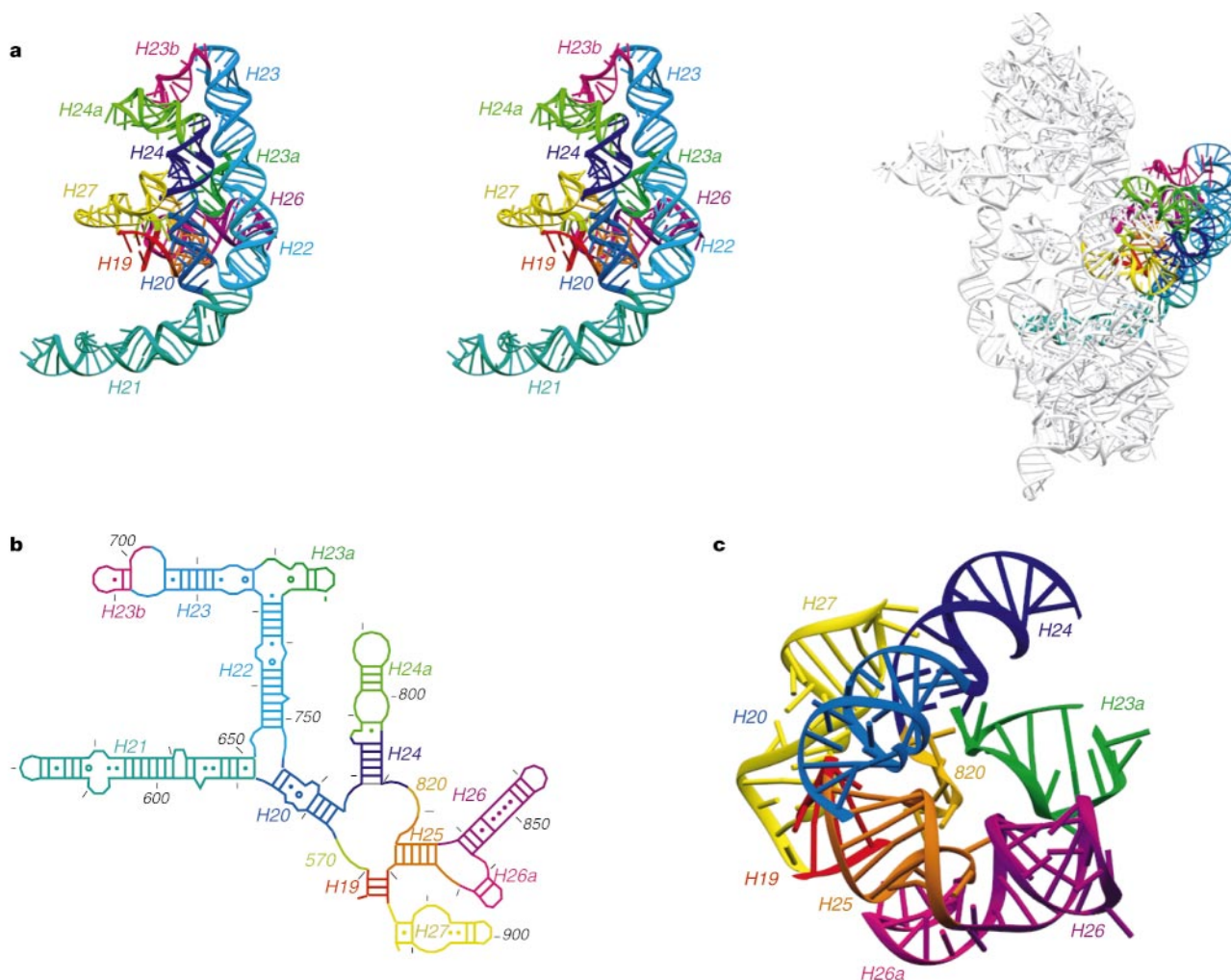


Figure 5 Structure of the central domain of 16S RNA. **a**, Stereo view of the domain with secondary structure diagram and inset showing its location in the 30S. **b**, Secondary

structure diagram for the central domain. **c**, Central portion of the domain, rotated relative to the other domains for clarity.

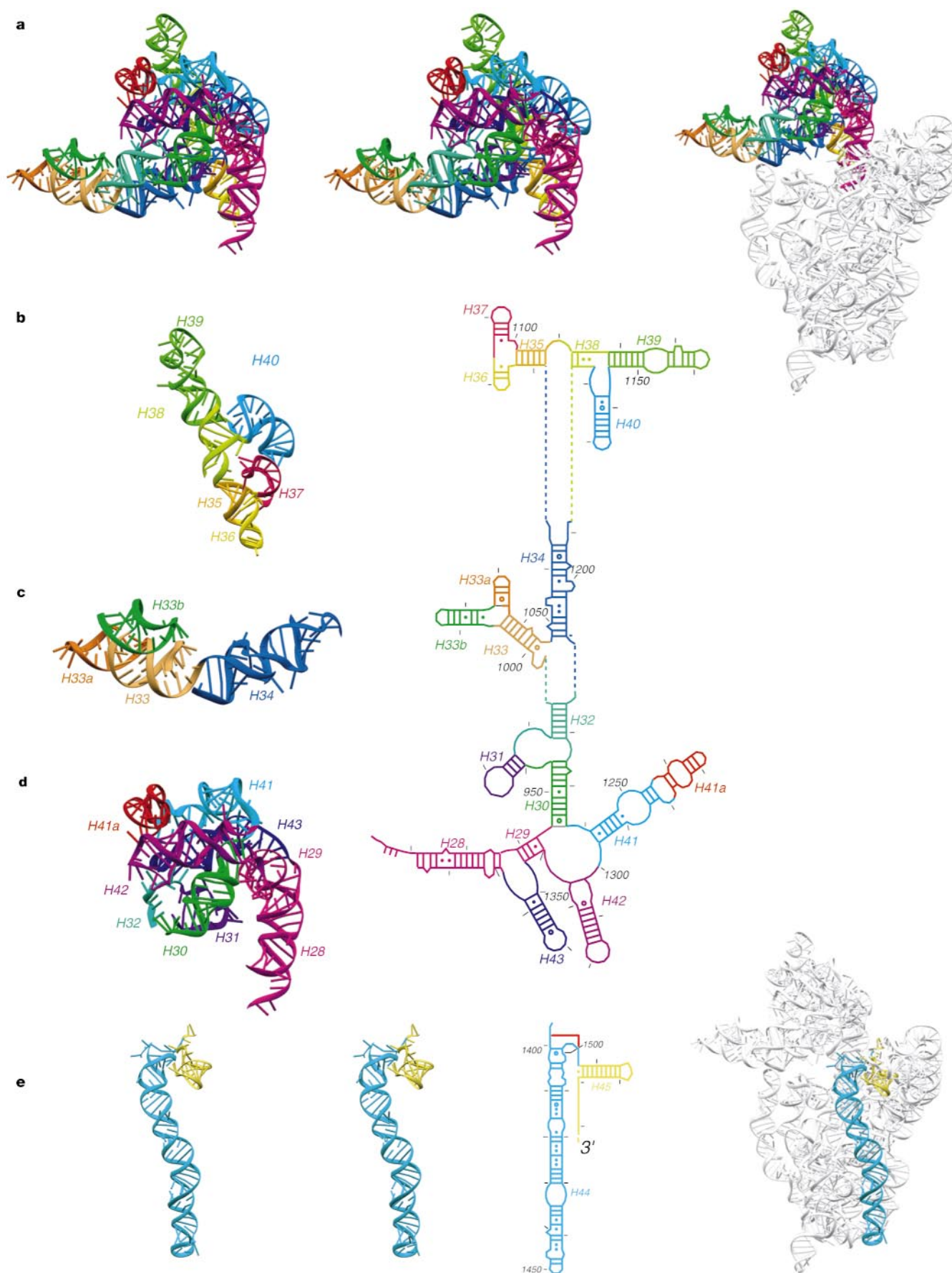


Figure 6 Structure of the 3' major and 3' minor domains of 16S RNA. **a**, Stereo view of the 3' major domain with inset showing its location in the 30S. **b–d**, The upper, middle and lower parts of the 3' major domain, with corresponding secondary structure

diagrams. **e**, Stereo view of the 3' minor domain, with secondary structure diagram and inset showing its location in the 30S.

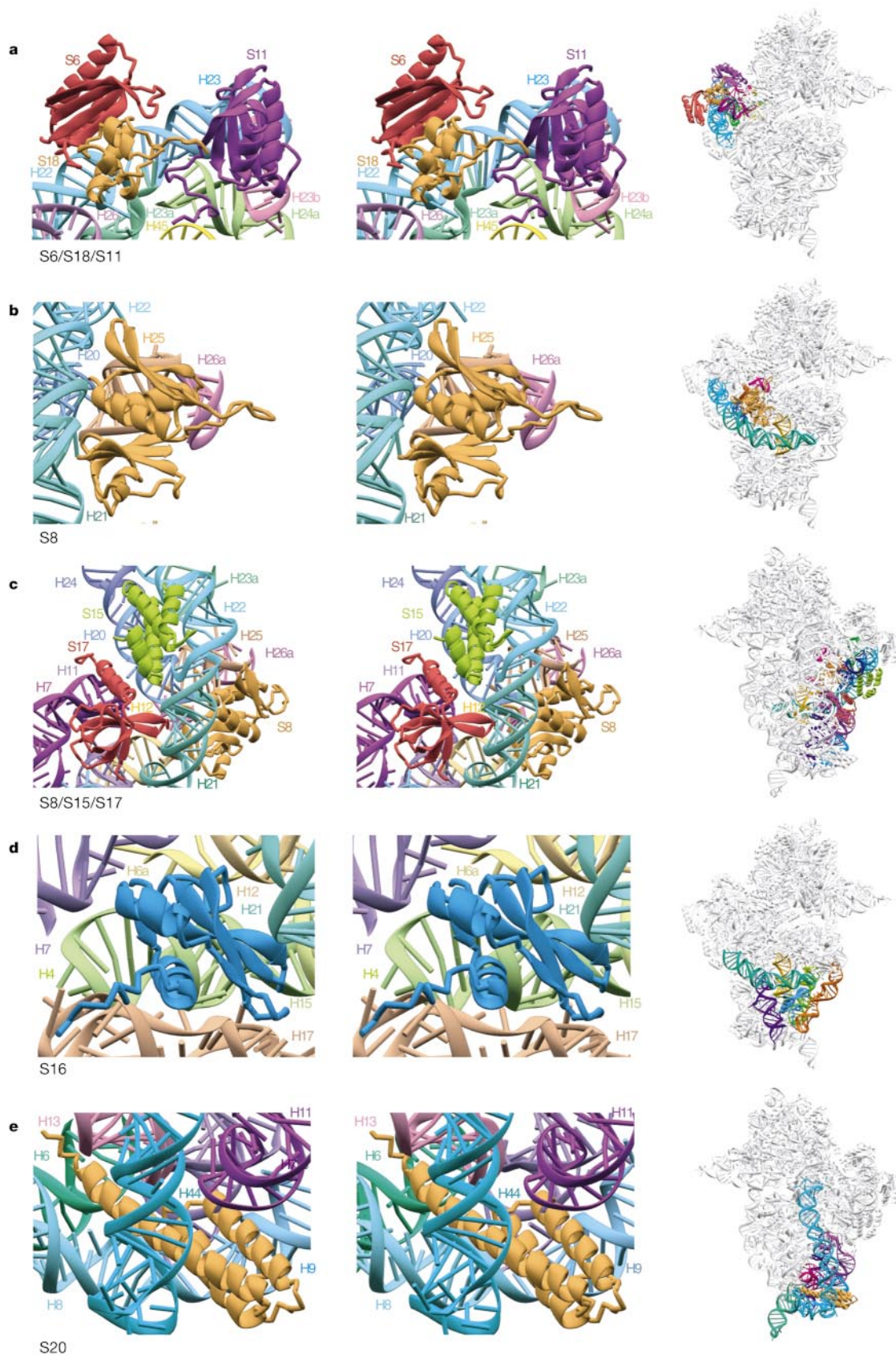


Figure 7 Proteins from the central and 5' domains. The colouring for the various RNA elements in this and subsequent figures is the same as in Figs 3–6. **a**, The S6–S18–S11

complex; **b**, S8; **c**, the S15–S17–S8 complex; **d**, S16; **e**, S20. In Figs 7–9 the stereo views have been rotated relative to the insets for clarity.

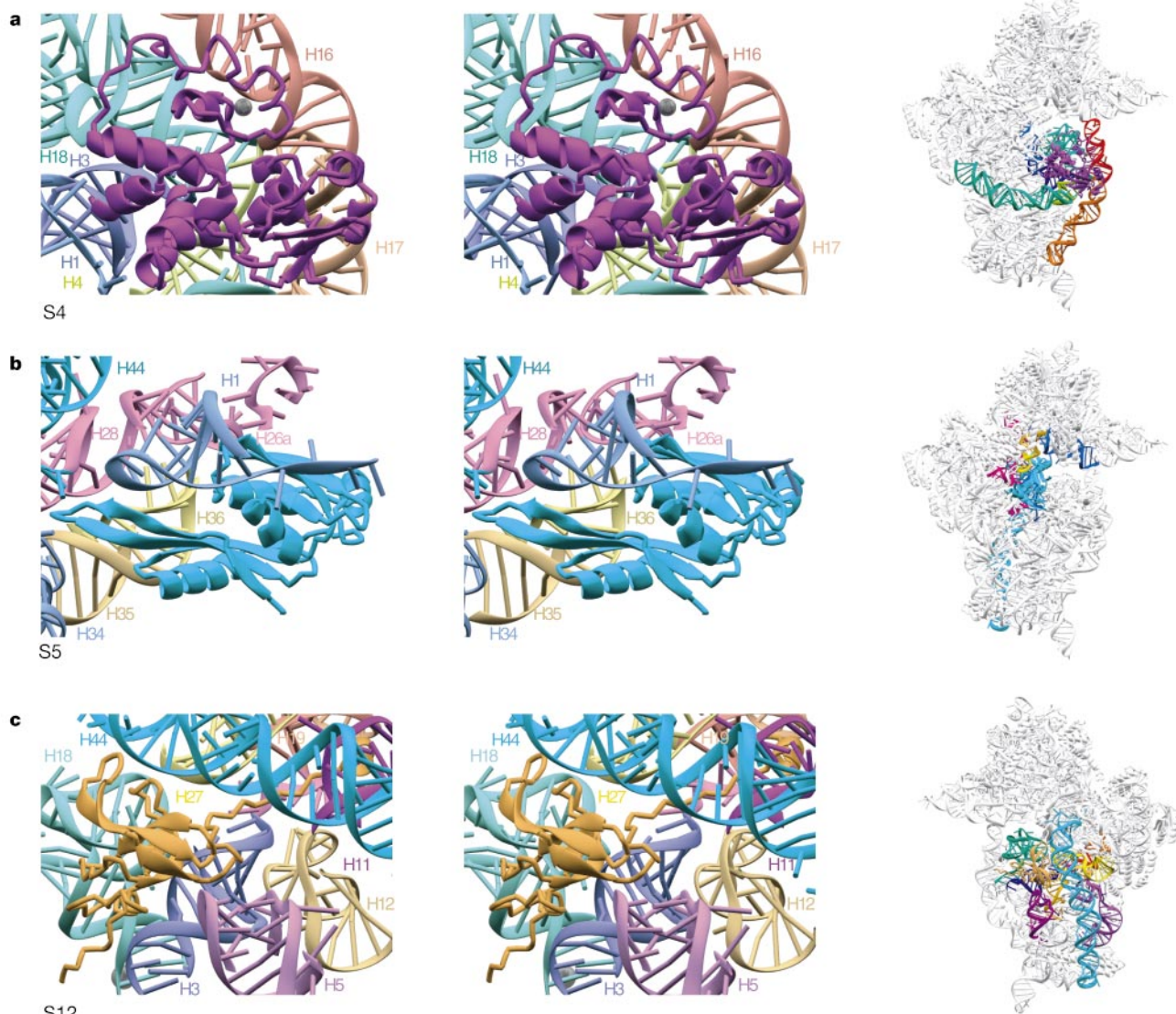
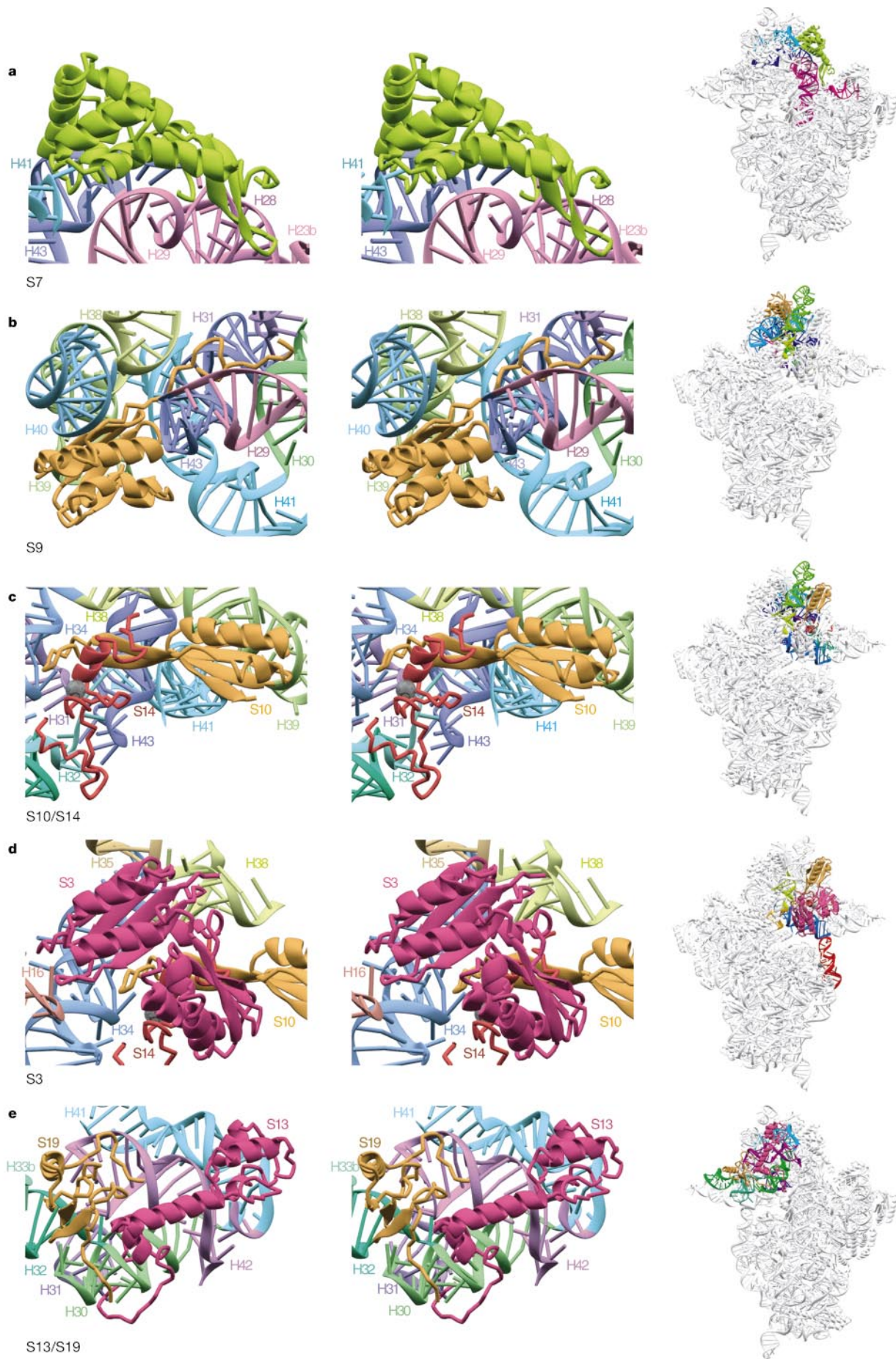


Figure 8 Proteins near the functional centre of the 30S. **a**, S4, with the Zn ion shown as a sphere; **b**, S5; **c**, S12.

Table 2 Summary of 30S protein structures

Protein	Figure no.	No. of residues (with N-terminal Met)	Visible residues	Previously known structures	No. of domains	Secondary structure in domains	RNA contacts	Protein interaction	Unusual features
S2	9f	256	V7–S235	No	2	$\alpha_2, \alpha_1\beta_5\alpha_3$	H26, H35, H36, H37, H40	None	Extended α -hairpin
S3	9d	239	G2–Q240	No	2	$\alpha_2\beta_3, \alpha_2\beta_4$	H16, H34, H35, H38, 530	S10, S14	N-terminal tail
S4	8a	209	G2–V207	Partial (refs 38, 39)	3	Zn finger, $\alpha_4, \alpha_3\beta_4$	H1, H3, H4, H16, H17, H18, H21	S5	N-terminal Zn finger
S5	8b	162	D5–G154	Yes (ref. 40)	2	$\alpha_1\beta_3, \alpha_2\beta_4$	H1, H2, H26a, H28, H34, H35, H36, H44, 560	S4, S8	extended β -hairpin
S6	7a	101	M1–A101	Yes (ref. 41)	1	$\alpha_2\beta_4$	H22, H23	S18	C-terminal tail
S7	9a	156	A2–W156	Yes (refs 42, 43)	1	$\alpha_6\beta_2$	H23b, H28, H29, H41, H43	S9, S11	extended β -hairpin
S8	7b,c	138	M1–W138	Yes (refs 44, 45)	2	$\beta_2\alpha_3, \alpha_1\beta_3$	H12, H20, H21, H22, H25, H26a	S5, S12, S17	
S9	9b	128	E2–R128	No	1	$\beta_3\beta_4$	H29, H30, H31, H38, H39, H40, H41, H43	S7	Long C-terminal tail
S10	9c, d	105	K3–T100	No	1	$\alpha_2\beta_4$	H31, H34, H38, H39, H41, H43	S3, S14	Long β -hairpin
S11	7a	129	K11–S129	No	1	$\alpha_2\beta_5$	H23, H23a, H23b, H24a, H45	S18, S7	Long N-terminal tail
S12	8c	135	P5–A128	No	1	$\alpha_1\beta_{1/5}$	H3, H5, H11, H12, H18, H19, H25, H27, 560, H44	S8, S17	Long N-terminal tail + extended β -hairpin loops
S13	9e	126	A2–K126	No	1	α_3	H30, H31, H41, H42	S19	Long C-terminal tail
S14	9c, d	61	A2–W61	No	1	none	H31, H32, H34, H38, 980, H42, H43	S3, S10	Zn module mostly extended
S15	7c	89	P2–G89	Yes (refs 22, 23, 46, 47)	1	α_4	H20, H22, H23a, H24	None	
S16	7d	88	M1–E83	Yes (ref. 48)	1	$\beta_1\alpha_2\beta_4$	H4, H6a, H7, H12, H15, H17, H21	None	C-terminal tail
S17	7c	105	P2–A105	Yes (ref. 49)	1	β_5	H7, H9, H11, H20, 560	S12	C-terminal helix + β -hairpin loops
S18	7a	88	P16–K88	Yes (ref. 23)	1	α_4	H22, H23, H23a, H26	S6; β -Strand extends S11 sheet	Registry and loop different from ref. 23
S19	9e	93	P2–R81	Yes (ref. 50)	1	$\alpha_1\beta_3$	H30, H32, H33b, H42	S13	
S20	7e	105	R8–A106	No	1	α_3	H6, H7, H8, H11, H13, H44	None	
Thx	9g	27	G2–K25	No	None	α_1	H30, H41, H41a, H42, H43	None	Mostly extended



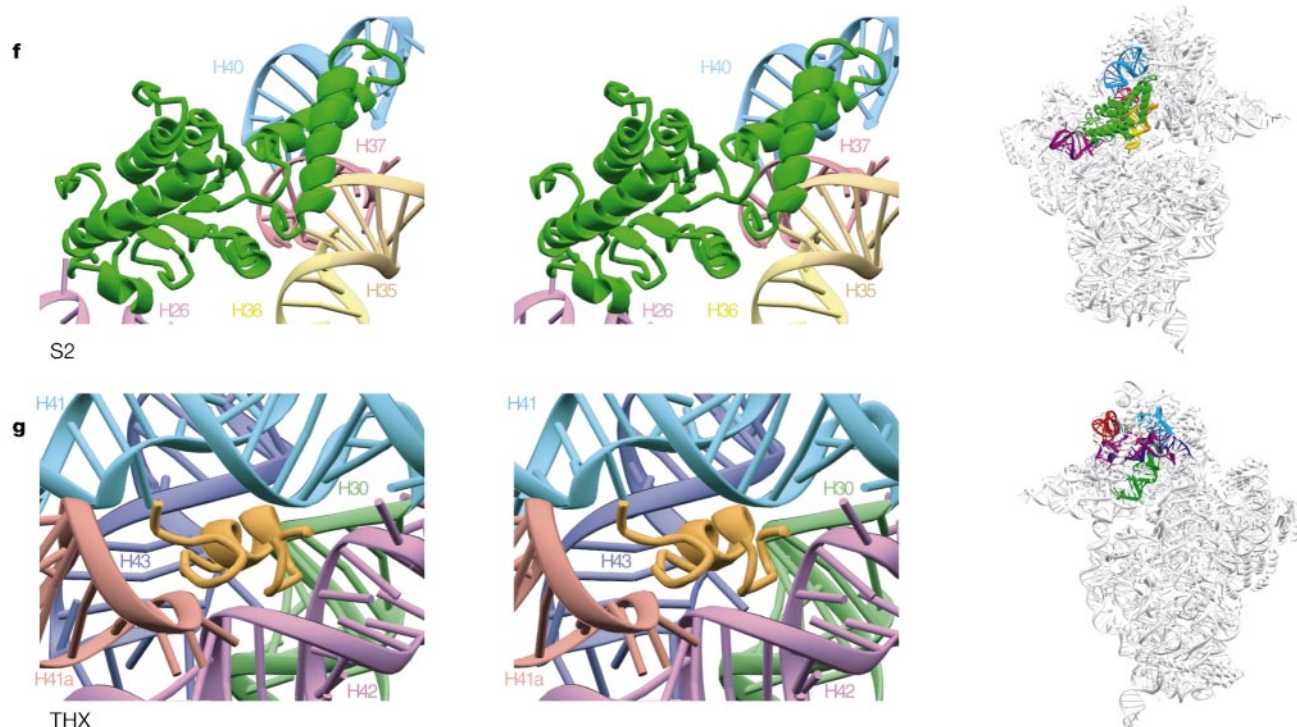


Figure 9 Proteins from the head. **a**, S7; **b**, S9; **c**, the S10–S14 complex, with the Zn ion in S14 shown as a sphere; **d**, S3 on top of S10 and S14; **e**, the S13–S19 complex; **f**, S2; **g**, Thx.

helices at the subunit interface (Fig. 6e). H44 is the longest single helix in the subunit, and stretches from the bottom of the head to the bottom of the body. It projects prominently from the body for interaction with the 50S subunit. H45 is roughly perpendicular to H44, with its conserved GGAA hairpin loop packed against H44 and available for interaction with the large subunit.

The 30S proteins and their interaction with 16S RNA

The structures of the proteins S2–S20 and Thx are summarized in Table 2 and shown with surrounding RNA and proteins in Figs 7–9. The structures of proteins solved in isolation (see refs in Table 2) were very useful in initial interpretation of the map³. In general, previous biochemical data on hydroxyl-radical footprinting²⁴ and ultraviolet-induced crosslinking (summarized in ref. 25 and http://www.mpimg-berlin-dahlem.mpg.de/~ag_ribo/ag_brimacombe/drc) agree well with the structure, and were useful as a guide to interpreting the fold at lower resolution. Finally, the structure is in good agreement with the neutron map²⁶, with S13, S14, S16, S19 and S20 being outliers. Of these, only S20 is located in a completely different region of the 30S from that predicted by the neutron map.

The proteins generally contain of one or more globular domains. The same types of folds are frequently found in different proteins, such as the β -barrel in S12 and S17, and α -helices packed against a β -sheet as observed in S3, S10, S6 and S11. It is interesting that these domains often interact with RNA in very different ways. The β -sheet of S11 packs flat against the minor groove of RNA, but in S6 the edge of the sheet interacts with RNA. In S3, one of the α/β -domains makes no contact with RNA at all.

In addition to a globular domain nearly all the proteins contain long extensions. These can be helical such as the α -hairpin in S2 or the carboxy-terminal helices in S13; long β -hairpins or loops protruding from S10 or S17; or extended carboxy- or amino-terminal tails seen in proteins S4, S9, S11, S12, S13 and S19. These extensions make intimate contact with RNA, and were

generally not visible in isolated structures because they were disordered in the absence of RNA. The extensions reach far into the surrounding RNA, and allow single proteins to contact several RNA elements, which is probably important for the stabilization of the RNA tertiary fold. The extensions are particularly well suited for this task because they are narrow, allowing close approach of different RNA elements, and they have the basic residues required to neutralize the charge repulsion of the RNA backbone. An extension of this principle is seen in Thx. This small peptide fits into a cavity between multiple RNA elements in the top of the head, and its positive charge stabilizes the organization of these elements.

Proteins are often found bound to junctions between helices. S4 binds to the 5-way junction between H3, H4, H16, H17, H18 in the 5' domain, and S7 binds tightly to the junction of H29, H30, H41, H42 in the 3' major domain. Both proteins are important in early stages of 30S assembly of the body and head respectively²⁷. Similarly, proteins S8, S15 and S17 bind to the three-way junction formed by H20, H21 and H22, and are important for the assembly of the central domain^{23,28}. Thus, protein binding to helical junctions is important for initiating the correct tertiary fold of RNA. Many of the proteins also contact RNA elements in different domains, helping to organize the overall structure of the 30S. For example, S17 contacts H7 and H11 in the 5' domain, and H21 in the central domain, while S20 mediates the contact between H44 in the 3' minor domain with the 5' domain at the base of the body. Much of the contact between the head and body is mediated by proteins, such as S2 and S5. In general the structure rationalizes biochemical studies on 30S assembly^{28,29}, although there are also discrepancies. For example, it is unlikely that the binding of S20 in the lower part of the 30S could influence the binding of S13 in the head. The implications of the structure for ribosome assembly will be published elsewhere.

In addition to the intimate contacts with RNA, a number of protein–protein interactions are seen. S3, S10 and S14 form a tight cluster held together by hydrophobic interactions. They are wedged

into a V-shaped gap between two RNA domains in the head, thus helping to stabilize its structure. In contrast, other protein clusters such as S4–S5–S8 are held together by predominantly electrostatic and hydrogen-bonding interactions, and the S4–S5 interface may undergo some rearrangement during 30S function, as discussed in the accompanying paper¹⁰. The globular domain of S12 on the interface side is connected by a long extension which snakes through the body of the RNA, and then folds into a short helix which makes contacts with S8 and S17 on the back side. A similar (although less extensive) interaction is observed between the C-terminal tail of S9 and the proteins S14 and S10 on the other side of the head. The tethering together of proteins that are on opposite sides of an RNA region may help hold the RNA structure together.

The proteins contain a number of special features that are probably required for the stability of *Thermus* 30S at high temperatures. An obvious example is the extra C-terminal helix in S17, which increases the RNA contacts made by this protein. Other examples are the Zn-binding motifs found in S14 and the N terminus of S4. The Zn-binding cysteine residues are not conserved, but the surrounding residues are, indicating that metal-ion coordination is used to give extra stability to a domain that is held together by hydrophobic contacts in mesophiles such as *E. coli*.

Conclusions

This high-resolution structure of the 30S ribosomal subunit is significant for several reasons. It will allow the rationalization in structural terms of four decades of biochemical efforts to elucidate the mechanism of protein synthesis. As a first step, functional insights from the structure and a description of 30S interactions with antibiotics are presented in the accompanying paper¹⁰. The structure will facilitate the design of testable models for various aspects of ribosome function, and it will also be a basis for the interpretation of lower resolution electron-microscopic or X-ray crystallographic maps of ribosomes in different functional states. The structure of the 30S also greatly increases our current database of RNA structure and protein–RNA interactions, and may provide general rules and improve the accuracy of prediction methods for RNA tertiary structure. □

Methods

Crystallization of the 30S subunit

We obtained crystals by a straightforward optimization of Trakhanov *et al.*⁴ with respect to pH, and concentrations of Mg²⁺ ions and MPD. The final conditions were 250 mM KCl, 75 mM NH₄Cl, 25 mM MgCl₂ and 6 mM 2-mercaptoethanol in 0.1 M potassium cacodylate or 0.1 M MES (2-*N*-morpholino-ethanesulphonic acid) at pH 6.5 with 13–17% MPD as the precipitant. We noticed initially that the 30S crystals completely lacked ribosomal protein S1 (J. L. C. May and V.R., unpublished results), so we removed S1 selectively from the 30S before crystallization, which improved both size and reproducibility. The crystals grew to the maximum size in about 6 weeks at 4 °C. The largest crystals, which were required for high-resolution data collection, grew to a size of 80–100 × 80–100 × 200–300 μm. The activity of redissolved crystals in poly(U)-directed protein synthesis was comparable to that of freshly isolated 30S subunits (data not shown).

Data collection

Crystals were transferred to 26% MPD by vapour diffusion in two steps over a period of six days. All solutions (except for those containing osmium hexamine or osmium pentamine) also contained 1 mM cobalt hexamine in the cryoprotectant. Crystals were flash-cooled by plunging into liquid nitrogen, and data collection was done in a cryostream at 90–100 K. A large fraction of crystals was screened at beamlines 9.6 or 14.1 at the SRS at Daresbury Laboratories, using two short exposures at least 40 degrees apart. These crystals were then analysed for diffraction limits, cell dimensions and mosaic spread. Only crystals of similar cell dimensions and with reasonable mosaic spread were used for data collection.

Potential derivatives were screened on beamlines X25 at the NLSL at Brookhaven National Laboratory and BM-14 at the ESRF (Grenoble). Data to about 4.5 Å were obtained from X25. High-resolution data were collected at SBC 19ID at the APS in Argonne National Laboratory, and ID14-4 at the ESRF. In all cases, derivative data were collected at the peak of the fluorescence at the LIII edge to maximize anomalous differences. At X25 and SBC 19ID, the kappa goniostat was used to rotate precisely about a mirror plane so that small anomalous differences could be measured accurately despite radiation decay and the use of multiple crystals. Each crystal typically yielded 3–10 degrees of data. Data were integrated and scaled using HKL-2000 (ref. 30).

Structure determination

Previously determined phases at 5.5 Å (ref. 3) were used to locate heavy atom sites using anomalous difference Fourier maps. Initially, these sites were used for phasing to 3.35 Å using the program SOLVE³¹, followed by density modification with SOLOMON³², using the procedure implemented in SHARP³³. Optimization of the various parameters in the procedure was required to obtain interpretable maps. The RNA and some of the proteins were built using the SOLVE maps. The sequence of *Thermus thermophilus* 16S RNA¹¹ and both previously available and unpublished (Göttingen *Thermus* Genome Sequencing Project) sequences for the proteins were used to build the structure. Improved maps were obtained by calculating experimental phases using SHARP³³ followed by density modification and phase extension to 3.05 Å with SOLOMON³² and DM³⁴. The improved maps allowed us to build all the remaining ordered parts of the structure. The model was built using O (ref. 35), and refined using the program CNS³⁶. Maximum likelihood refinement was used, initially with both amplitudes and experimental phase probability distributions to 3.35 Å, and subsequently with amplitudes to 3.05 Å.

Details of the experimental protocols will be published elsewhere.

Received 14 July; accepted 10 August 2000.

- Garrett, R. A. *et al.* (eds) *The Ribosome. Structure, Function, Antibiotics and Cellular Interactions* (ASM, Washington DC, 2000).
- von Böhlen, K. *et al.* Characterization and preliminary attempts for derivatization of crystals of large ribosomal subunits from *Haloarcula marismortui* diffracting to 3 Å resolution. *J. Mol. Biol.* **222**, 11–15 (1991).
- Clemons, W. M. *et al.* Structure of a bacterial 30S ribosomal subunit at 5.5 Å resolution. *Nature* **400**, 833–840 (1999).
- Trakhanov, S. D. Crystallization of 70S ribosomes and 30S ribosomal subunits from *Thermus thermophilus*. *FEBS Lett.* **220**, 319–322 (1987).
- Glotz, C. *et al.* Three-dimensional crystals of ribosomes and their subunits from eu- and archaeobacteria. *Biochem. Int.* **15**, 953–960 (1987).
- Yonath, A. *et al.* Characterization of crystals of small ribosomal subunits. *J. Mol. Biol.* **203**, 831–834 (1988).
- Yusupov, M. M., Tischenko, S. V., Trakhanov, S. D., Ryazantsev, S. N. & Garber, M. B. A new crystalline form of 30S ribosomal subunits from *Thermus thermophilus*. *FEBS Lett.* **238**, 113–115 (1988).
- Yonath, A. Crystallographic studies on the ribosome, a large macromolecular assembly exhibiting severe nonisomorphism, extreme beam sensitivity and no internal symmetry. *Acta Crystallogr. A* **54**, 945–955 (1998).
- Tocij, A. *et al.* The small ribosomal subunit from *Thermus thermophilus* at 4.5 Å resolution: pattern fittings and the identification of a functional site. *Proc. Natl Acad. Sci. USA* **96**, 14252–14257 (1999).
- Carter, A. P. *et al.* Functional insights from the structure of the 30S ribosomal subunit and its interaction with antibiotics. *Nature* **407**, 340–348 (2000).
- Hartmann, R. K. & Erdmann, V. A. *Thermus thermophilus* 16S rRNA is transcribed from an isolated transcription unit. *J. Bacteriol.* **171**, 2933–2941 (1989).
- Choli, T., Franceschi, F., Yonath, A. & Wittmann-Liebold, B. Isolation and characterization of a new ribosomal protein from the thermophilic eubacteria, *Thermus thermophilus*, *T. aquaticus* and *T. flavus*. *Biol. Chem. Hoppe Seyler* **374**, 377–383 (1993).
- Mueller, F. & Brimacombe, R. A new model for the three-dimensional folding of *Escherichia coli* 16S ribosomal RNA. I. Fitting the RNA to a 3D electron microscopic map at 20 Å. *J. Mol. Biol.* **271**, 524–544 (1997).
- Gabashvili, I. S., Agrawal, R. K., Grassucci, R. & Frank, J. Structure and structural variations of the *Escherichia coli* 30S ribosomal subunit as revealed by three-dimensional cryo-electron microscopy. *J. Mol. Biol.* **286**, 1285–1291 (1999).
- Gabashvili, I. S. *et al.* Solution structure of the *E. coli* 70S ribosome at 11.5 Å resolution. *Cell* **100**, 537–549 (2000).
- Stern, S., Weiser, B. & Noller, H. F. Model for the three-dimensional folding of 16S ribosomal RNA. *J. Mol. Biol.* **204**, 447–481 (1988).
- Malhotra, A. & Harvey, S. C. A quantitative model of the *Escherichia coli* 16S RNA in the 30S ribosomal subunit. *J. Mol. Biol.* **240**, 308–340 (1994).
- Ramakrishnan, V. Distribution of protein and RNA in the 30S ribosomal subunit. *Science* **231**, 1562–1564 (1986).
- Wimberly, B., Varani, G. & Tinoco, I. The conformation of loop E of eukaryotic 5S ribosomal RNA. *Biochemistry* **32**, 1078–1087 (1993).
- Lodmell, J. S. & Dahlberg, A. E. A conformational switch in *Escherichia coli* 16S ribosomal RNA during decoding of messenger RNA. *Science* **277**, 1262–1267 (1997).
- Gutell, R. R. in *Ribosomal RNA Structure, Evolution, Processing and Function in Protein Biosynthesis* (eds Dahlberg, A. E. & Zimmermann, R. A.) 111–128 (CRC, Boca Raton, 1996).
- Nikulin, A. *et al.* Crystal structure of the S15+ rRNA complex. *Nature Struct. Biol.* **7**, 273–277 (2000).
- Agalarov, S. C., Sridhar Prasad, G., Funke, P. M., Stout, C. D. & Williamson, J. R. Structure of the S15₅₆S18-rRNA complex: assembly of the 30S ribosome central domain. *Science* **288**, 107–113 (2000).
- Powers, T. & Noller, H. F. Hydroxyl radical footprinting of ribosomal proteins on 16S rRNA. *RNA* **1**, 194–209 (1995).
- Mueller, F. & Brimacombe, R. A new model for the three-dimensional folding of *Escherichia coli* 16S ribosomal RNA. II. The RNA-protein interaction data. *J. Mol. Biol.* **271**, 545–565 (1997).
- Capel, M. S. *et al.* A complete mapping of the proteins in the small ribosomal subunit of *Escherichia coli*. *Science* **238**, 1403–1406 (1987).
- Nowotny, V. & Nierhaus, K. H. Assembly of the 30S subunit from *Escherichia coli* ribosomes occurs via two assembly domains which are initiated by S4 and S7. *Biochemistry* **27**, 7051–7055 (1988).
- Stern, S., Powers, T., Changchien, L.-M. & Noller, H. F. RNA-protein interactions in 30S ribosomal subunits: folding and function of 16S rRNA. *Science* **244**, 783–790 (1989).
- Held, W. A., Ballou, B., Mizushima, S. & Nomura, M. Assembly mapping of 30S ribosomal proteins from *Escherichia coli*. Further studies. *J. Biol. Chem.* **249**, 3103–3111 (1974).
- Otwinowski, Z. & Minor, W. in *Methods in Enzymology* (eds Carter, C. W. & Sweet, R. M.) 307–325 (Academic, New York, 1997).
- Terwilliger, T. & Berendzen, J. Automated MAD and MIR structure determination. *Acta Crystallogr. D* **55**, 849–861 (1999).

32. Abrahams, J. P. Bias reduction in phase refinement by modified interference functions: introducing the gamma correction. *Acta Crystallogr. D* **53**, 371–376 (1997).
33. de la Fortelle, E. & Bricogne, G. in *Methods in Enzymology* (eds Carter, C. W. & Sweet, R. M.) 472–493 (Academic, New York, 1997).
34. Cowtan, K. & Main, P. Miscellaneous algorithms for density modification. *Acta Crystallogr. D* **54**, 487–493 (1998).
35. Jones, T. A. & Kjeldgaard, M. Electron-density map interpretation. *Methods Enzymol.* **277B**, 173–207 (1997).
36. Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
37. Carson, M. Ribbons 2.0. *J. Appl. Cryst.* **24**, 958–961 (1991).
38. Davies, C., Gerstner, R. B., Draper, D. E., Ramakrishnan, V. & White, S. W. The crystal structure of ribosomal protein S4 reveals a two-domain molecule with an extensive RNA-binding surface: one domain shows structural homology to the ETS DNA-binding motif. *EMBO J.* **17**, 4545–4558 (1998).
39. Markus, M. A., Gerstner, R. B., Draper, D. E. & Torchia, D. A. The solution structure of ribosomal protein S4 delta41 reveals two subdomains and a positively charged surface that may interact with RNA. *EMBO J.* **17**, 4559–4571 (1998).
40. Ramakrishnan, V. & White, S. W. Structure of ribosomal protein S5 reveals sites of interaction with 16S RNA. *Nature* **358**, 768–771 (1992).
41. Lindahl, M. *et al.* Crystal structure of the ribosomal protein S6 from *Thermus thermophilus*. *EMBO J.* **13**, 1249–1254 (1994).
42. Wimberly, B. T., White, S. W. & Ramakrishnan, V. The structure of ribosomal protein S7 at 1.9 Å resolution reveals a beta-hairpin motif that binds double-stranded nucleic acids. *Structure* **5**, 1187–1198 (1997).
43. Hosaka, H. *et al.* Ribosomal protein S7: a new RNA-binding motif with structural similarities to a DNA architectural factor. *Structure* **5**, 1199–1208 (1997).
44. Davies, C., Ramakrishnan, V. & White, S. W. Structural evidence for specific S8-RNA and S8-protein interactions within the 30S ribosomal subunit: ribosomal protein S8 from *Bacillus stearothermophilus* at 1.9 Å resolution. *Structure* **4**, 1093–1104 (1996).
45. Nevskaya, N. *et al.* Crystal structure of ribosomal protein S8 from *Thermus thermophilus* reveals a high degree of structural conservation of a specific RNA binding site. *J. Mol. Biol.* **279**, 233–244 (1998).
46. Berglund, H., Rak, A., Serganov, A., Garber, M. & Härd, T. Solution structure of the ribosomal RNA binding protein S15 from *Thermus thermophilus*. *Nature Struct. Biol.* **4**, 20–23 (1997).
47. Clemons, W. M., Davies, C., White, S. W. & Ramakrishnan, V. Conformational variability of the N-terminal helix in the structure of ribosomal protein S15. *Structure* **6**, 429–438 (1998).
48. Allard, P. Another piece of the ribosome: Solution structure of S16 and its location in the 30S subunit. *Structure* **8**, 875–882 (2000).
49. Golden, B. L., Hoffman, D. W., Ramakrishnan, V. & White, S. W. Ribosomal protein S17: characterization of the three-dimensional structure by ¹H- and ¹⁵N-NMR. *Biochemistry* **32**, 12812–12820 (1993).
50. Helgstrand, M. *et al.* Solution structure of the ribosomal protein S19 from *Thermus thermophilus*. *J. Mol. Biol.* **292**, 1071–1081 (2000).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>).

Acknowledgements

This work was supported by the Medical Research Council (UK) and a US National Institutes of Health grant to V.R. and S. W. White. Beamlines at Argonne and Brookhaven were supported by the US Department of Energy. D.E.B. was supported by an EMBO long-term postdoctoral fellowship, and W.M.C. by an NIH predoctoral fellowship. We thank B. S. Brunschwig and M. H. Chou for gifts of osmium hexammine and osmium bipyridine; T. Terwilliger for help with phasing using SOLVE; T. A. Leaf-Jones for providing us a version of O with RNA tools; and our colleagues at the LMB for their advice and encouragement. We are indebted to A. Joachimiak, S. L. Ginell, R. Ravelli, S. McSweeney, G. Leonard, A. Thompson, H. Lewis, L. Berman, M. Papiz, S. Girdwood and M. MacDonald for help and advice on synchrotron beamlines.

Correspondence and requests for materials should be addressed to V.R. (e-mail: ramak@mbc.cam.ac.uk). Coordinates have been deposited in the Protein Data Bank, accession number 1FJF. Coordinates of individual components will be made available on <http://alf1.mrc-mb.cam.ac.uk/~ramak/30S>.

Functional insights from the structure of the 30S ribosomal subunit and its interactions with antibiotics

Andrew P. Carter^{*†‡}, William M. Clemons^{*†‡}, Ditlev E. Brodersen^{*†‡}, Robert J. Morgan-Warren[†], Brian T. Wimberly[†] & V. Ramakrishnan[†]

[†] MRC Laboratory of Molecular Biology, Hills Road, Cambridge, UK

[‡] Department of Biochemistry, University of Utah School of Medicine, Salt Lake City, Utah 84132, USA

* These authors contributed equally to this work.

The 30S ribosomal subunit has two primary functions in protein synthesis. It discriminates against aminoacyl transfer RNAs that do not match the codon of messenger RNA, thereby ensuring accuracy in translation of the genetic message in a process called decoding. Also, it works with the 50S subunit to move the tRNAs and associated mRNA by precisely one codon, in a process called translocation. Here we describe the functional implications of the high-resolution 30S crystal structure presented in the accompanying paper, and infer details of the interactions between the 30S subunit and its tRNA and mRNA ligands. We also describe the crystal structure of the 30S subunit complexed with the antibiotics paromomycin, streptomycin and spectinomycin, which interfere with decoding and translocation. This work reveals the structural basis for the action of these antibiotics, and leads to a model for the role of the universally conserved 16S RNA residues A1492 and A1493 in the decoding process.

The ribosome is a large, template-directed enzyme with sophisticated proof-reading capabilities. Its ligands are tRNA and mRNA. There are three tRNA-binding sites, designated A (aminoacyl), P (peptidyl) and E (exit), after their respective tRNA substrates (reviewed in ref. 1). The anticodon stem-loops (ASL) of the A- and P-site tRNAs bind to the 30S where they are base paired with adjacent codons on mRNA. The E-site tRNA is bound in a similar orientation, but whether the E-site tRNA is base-paired to the E-site mRNA codon or not is a matter of controversy^{2,3}.

The decoding of mRNA into protein requires the correct recognition of each A-site codon by the anticodon of the corresponding aminoacyl-tRNA (aa-tRNA). The difference in binding energy between the codon and anticodon with correct (cognate) versus incorrect (non-cognate) tRNAs is too small to account for the high accuracy of translation⁴. Initially, a ternary complex of elongation factor Tu (EF-Tu), aminoacyl tRNA and GTP binds to the A-site. The high level of accuracy of translation by the ribosome is thought to be a combination of the initial codon-anticodon interaction and a proof-reading step occurring after GTP hydrolysis and release of EF-Tu^{5,6}. The proof-reading step is considered to be especially important for discriminating cognate codon-anticodon interactions from energetically similar near-cognate ones. The mechanism by which proof-reading occurs remains unclear, although a number of models have been proposed^{7,8}. There are also models for decoding that do not invoke a separate proof-reading step⁹.

It has been suggested that the 30S switches between at least two distinct states during translation, and that stabilizing one state over the other can affect accuracy^{10,11}. One genetically well characterized conformational switch in the 30S subunit is the helix 27 accuracy switch¹¹. Genetic and biochemical data support a model in which this switch may be part of a larger-scale conformational change that occurs between initial selection and proof-reading of the A-site tRNA, or it may have a role in translocation.

Until recently, there has been a large disparity between the high resolution of the genetic and biochemical data that define the RNA components of the active sites of the 30S subunit, and the relatively low resolution of the three-dimensional structures of ribosomes available. Here we analyse the high-resolution structure of the 30S

from *Thermus thermophilus* reported in the accompanying paper¹², and identify the molecular details of the interaction of mRNA and tRNA with the A-, P- and E-sites of the 30S. We also discuss the structural implications of the switch involving helix 27.

Both subunits are targets for antibiotics. By interfering with various aspects of their function, these antibiotics help to shed light on the mechanisms involved in translation. Although antibiotics were characterized several decades ago, their molecular mechanisms were not clear in the absence of a high-resolution structure. We describe here the crystal structure to 3 Å resolution of 30S from *Thermus thermophilus* complexed with streptomycin and paromomycin (which reduce translational accuracy), and spectinomycin (which interferes with translocation). Difference Fourier maps calculated after the refinement of our atomic model of the 30S against these data reveal the bound antibiotic molecules *in situ*. The structure allows us to rationalize much of the biochemical and genetic data on these antibiotics and propose models for their mode of action. In addition, the effect of paromomycin on the universally conserved bases A1492 and A1493 indicates a model for how they participate in the decoding process.

Results

A single round of refinement against the native 30S coordinates resulted in a model with R/R_{free} of 0.231/0.266. During refinement, the 30S structure changed significantly only in the vicinity of the paromomycin-binding site. Elsewhere, the phosphate r.m.s. deviation between the antibiotic-bound and free 30S structures fluctuated by about 0.45 Å, which is comparable to the coordinate error in the original model. Electron density for the antibiotics was visible in both difference Fourier maps and σ -weighted $2mF_o - DF_c$ maps. At 3 Å resolution, it was straightforward to position each antibiotic in an unambiguous orientation in the density, so that precise conclusions could be made about the location and interactions of each antibiotic in the 30S subunit. A second round of refinement using a model that included the antibiotics reduced the R/R_{free} to 0.224/0.258. Crystallographic statistics are shown in Table 1 of the Supplementary Information.

In the following discussion, we analyse the 30S structure with

respect to function. As described in the accompanying paper, we use the standard H1–H45 nomenclature¹³ to describe the helical elements in 16S RNA and the numbering used is for the *Escherichia coli* 16S RNA sequence¹².

Mimics of P-site tRNA and mRNA

Functionally, the most important parts of the 30S subunit are its substrate-binding A-, P- and E-sites. In our 30S crystal structure¹², the P-site is occupied by the tip of the spur stem-loop (H6) from a symmetry-related 30S subunit. The spur thus appears to mimic the P-site tRNA anticodon stem-loop. A superposition of the 7.8 Å 70S structure including mRNA and tRNA¹⁴ onto our 30S structure using the common elements H27 and H44 gave a phosphate r.m.s. deviation of only 2.3 Å after rejection of outliers. This superposition shows residues 27–43 of P-site tRNA almost exactly coinciding with residues 75–95 of the spur (Fig. 1a), indicating that it is a good mimic of the ASL. Our 30S maps also contain electron density for single-stranded RNA that mimics mRNA codons in the P- and E-sites, with base pairing between the P-site codon and residues that correspond to the anticodon of the spur. This mRNA mimic is of uncertain origin, but its location and sequence (UCU, inferred from the non Watson–Crick base-pairing geometries with the UUU of the spur) indicate that it corresponds to the 3' end of 16S RNA. The last nucleotide of our 16S model is A1534, which implies that four disordered nucleotides 1535–1538 would span the 7 Å distance to the first nucleotide of the E-site codon mimic. Alternatively, it is possible that the 3' end of 16S RNA has been cleaved somewhere between A1534 and U1539 before or during crystallization. Spur coordinates are available in the Supplementary Information.

Definition of the 30S A-, P- and E-sites

The 16S RNA-based superposition of the 70S tRNAs and mRNA onto the 30S structure also provides a basis for defining the A-, P- and E-sites on the 30S subunit (Fig. 1b–d). There are no serious clashes between the A-, P- and E-site tRNAs and the 30S subunit (Fig. 1b). The 30S A-, P- and E-site elements defined in this manner are shown in Fig. 1c and d. The most striking characteristic of the three binding sites is that they are all composed of RNA elements from at least two different domains (Fig. 1c). It is clear that independent motion of the head (3' domain) or platform (central domain) observed in low-resolution cryo-EM studies^{15,16} would have important consequences for movement of the tRNA–mRNA complex, as must occur during translocation.

The P-site. The spur contacts several discrete regions of 16S RNA, and surprisingly, two proteins (Fig. 2a and b). Most of the interaction surface is found on the minor groove of the spur stem and RNA nucleotides 1338–1341, 1229–1230 and small portions of the carboxy-terminal tails of proteins S13 and S9. There are seven hydrogen bonds from the minor groove of spur residues C91–C92 and G78 to the minor groove surface of G1338–A1339. Only one of these hydrogen bonds appears to be sequence specific. Both G1338 and A1339 have been implicated in P-site binding¹⁷. A contact from K126 of S9 appears to help to stabilize this minor-groove-to-minor-groove packing interaction. A second area of contact, nearly continuous with the first, is the interaction between the 16S 1229–1230 sugar–phosphate backbone and spur residues G77 and G78. This region of contact is extended by the C-terminal tail of S13. The other areas of contact are less extensive. One interaction is the stacking of U82 on C1400, which rationalizes

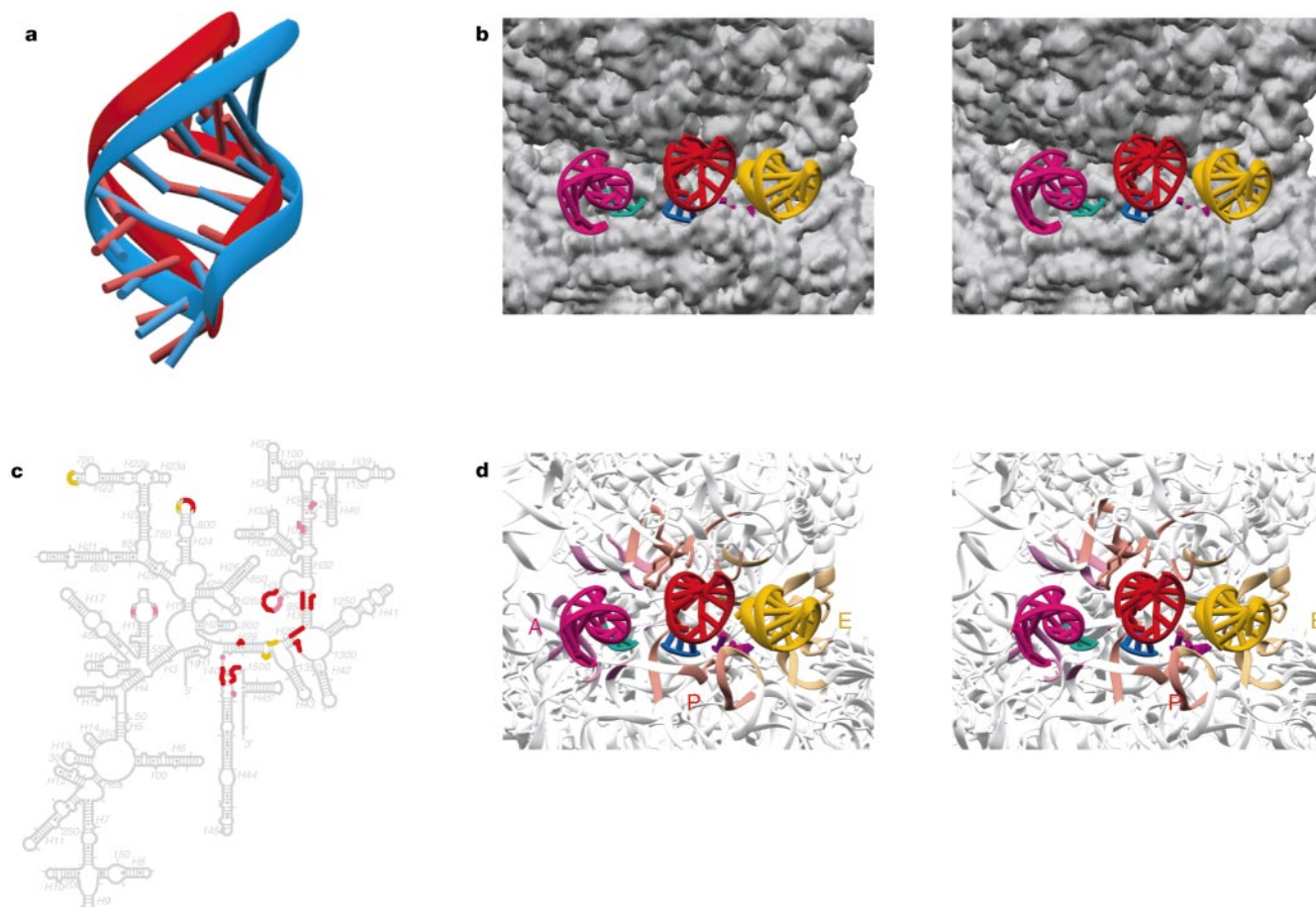


Figure 1 Overview of the 30S A-, P- and E-sites. **a**, Superposition of the 70S P-site tRNA anticodon stem-loop (ASL)¹⁴ with its mimic in the 30S structure, the spur stem-loop. **b**, Model of the A-, P- and E-site mRNA and tRNA ASLs on the 30S surface. **c**, Secondary

structure of 16S RNA. Magenta, A-site; red, P-site; yellow, E-site. **d**, Model of the A-, P- and E-site mRNA and tRNA ASLs with 30S components of the three sites coloured as in **c**. This and other figures were prepared with RIBBONS⁴⁶.

the ASL34–C1400 ultraviolet-induced crosslink¹⁸. The other region is a packing interaction between A790 and spur residues 88–89, with a single hydrogen bond present. Finally, if the spur anticodon–mRNA codon helix were a few ångströms wider, as it would be for a Watson–Crick-paired helix, it would make van der Waals contact with the base of G966, which has been implicated as part of the P-site by chemical modification and binding experiments^{17,19}.

The P-site codon threads through the major groove of the upper portion of helix 44, in a universally conserved region of 16S RNA (Fig. 2a, b and e). There is a tight turn between nucleotides –1 and +1, between the last E-site and the first P-site codon nucleotides. This turn is stabilized by a hydrogen bond to the N1/N2 groups of the conserved residue G926, a residue previously implicated in P-site binding¹⁹. Additional hydrogen bonds are observed between

the 2' OH of nucleotide +1 to the phosphate of C1498, and between the phosphate of nucleotide +2 and the 2' OH of C1498. The phosphate of +2 also stacks on the base of C1498. The phosphate of +3 is within hydrogen-bonding distance of two conserved cytidine N4 groups from C1402 and C1403. The +3 base also stacks on the sugar of C1400. Finally, it appears probable that there are several Mg²⁺ ions that may help stabilize the location of the P-site codon in the major groove of H44. One such ion appears to interact with the N7 atom of G1401, in agreement with biochemical data implicating this residue in P-site function^{17,19}. The contacts are also consistent with mRNA crosslinking data (summarized in ref. 20).

The A-site. The A-site is much wider and shallower than the P- or E-sites (Figs 1b and 2c), consistent with its much lower affinity for tRNA. The shallowness of the A-site may reflect the need to allow

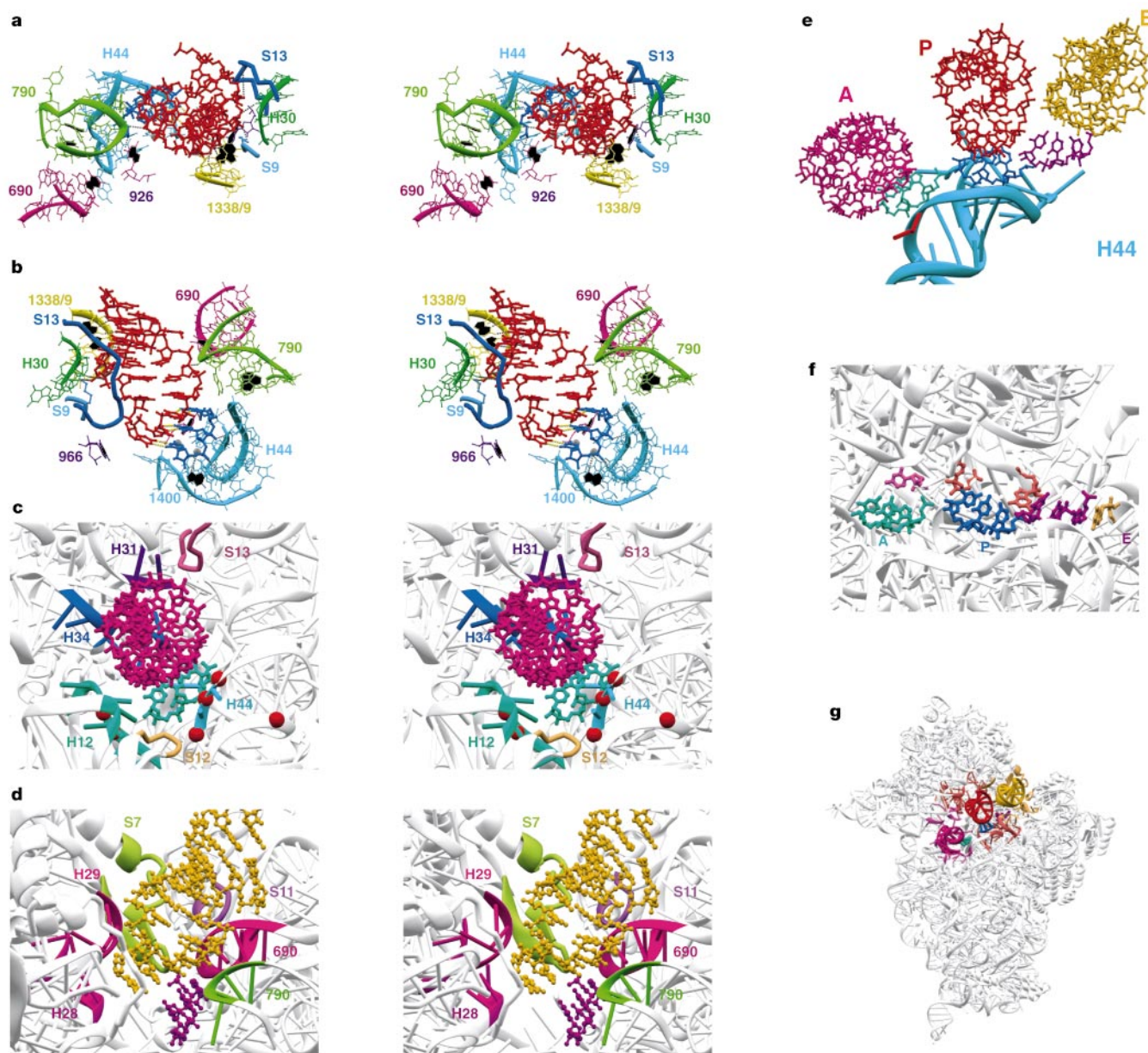


Figure 2 Detailed views of the 30S A-, P-, and E-sites with models of bound tRNA and mRNA ligands. **a, b**, Two roughly perpendicular views of the spur in the P-site. Dotted yellow lines indicate base pairing between the spur (red) and the codon mimic (dark blue). Dotted black lines indicate hydrogen bonds. RNA bases coloured black have been implicated in binding of P-site tRNA or mRNA (see text). **c**, Stereo view of the A-site, including the A-site ASL (magenta) and codon (green) modelled as in Fig. 1. Red balls

highlight RNA residues implicated in binding of A-site tRNA or mRNA by biochemical experiments (see text). **d**, Stereo view of the E-site, including the E-site codon from the 30S structure (purple) and the ASL (yellow) modelled as in Fig. 1. **e, f**, Two views of the mRNA model shown in **a–d** (see text). **g**, Inset showing the location of the A-, P- and E-sites on the 30S subunit.

rotation of the A-site codon–anticodon helix during or after GTP hydrolysis by EF-Tu. The RNA components of the A-site include portions of the 530 loop, H31 and H34 in the head, as well as residues 1492–1493 from the 3' minor domain (Fig. 2c), all of which have been implicated as elements of the decoding site (see articles in ref. 21). Interestingly, the electron density for A1492 and A1493 is not consistent with a single conformation for these residues in the absence of antibiotics. The only protein that appears even remotely likely to participate directly in decoding is S12, whose K47 loop is close to the A-site codon–anticodon helix. Other proteins are more distant: a poorly conserved part of the C-terminal tail of S13 lies between the A- and P-sites, and the conserved but disordered C-terminal 12 residues of S19 may be in approximately the right location to interact with the upper portion of the A-site ASL.

The E-site. Unlike the A- and P-sites, the E-site consists mostly of protein (Fig. 2d). Proteins S7 and S11 have a small interface that binds the minor groove of the E-site ASL. The position of the highly conserved β -hairpin of S7 indicates that it might help dissociate the E-site codon–anticodon. The RNA portion of the E-site makes relatively few interactions with the modelled E-site ASL, in accordance with the observed failure of the E-site tRNA to footprint 16S

RNA¹⁷. Several of the RNA residues footprinted by P-site tRNA, including G693, A794 and C795 are found to be in contact with the E-site codon¹⁷.

The helix 27 accuracy switch. In addition to the ligand-binding A-, P- and E- sites, the 30S subunit must also contain conformational switches that are important for its function. Only one of these, involving H27, has been well characterized genetically and biochemically¹¹. H27 has been proposed to have two alternative base-pairing schemes during translation: a ribosomal ambiguity (*ram*) or 'error-prone' form with nucleotides 885–887 paired to 910–912, and an alternative hyperaccurate or 'restrictive' form with nucleotides 888–890 paired to 910–912. The *ram* form features an S-turn motif in H27, which is present in our structure (green in Fig. 3a). It packs against the minor groove of H44, just below the decoding site (Fig. 3b). Switching to the restrictive form must disrupt the S-turn, and change the packing of H27 and H44 and perhaps the structure of the decoding site. Indeed, strains containing S12 restrictive mutations exhibit altered chemical reactivity of RNA residues at the H27–H44 interface (orange and magenta in Fig. 3b)¹⁰. However, H27 also packs against several other helices: the conserved hairpin loop at the end of H27 packs against the platform (H24), and there is a long-range base pair between H27 and the root

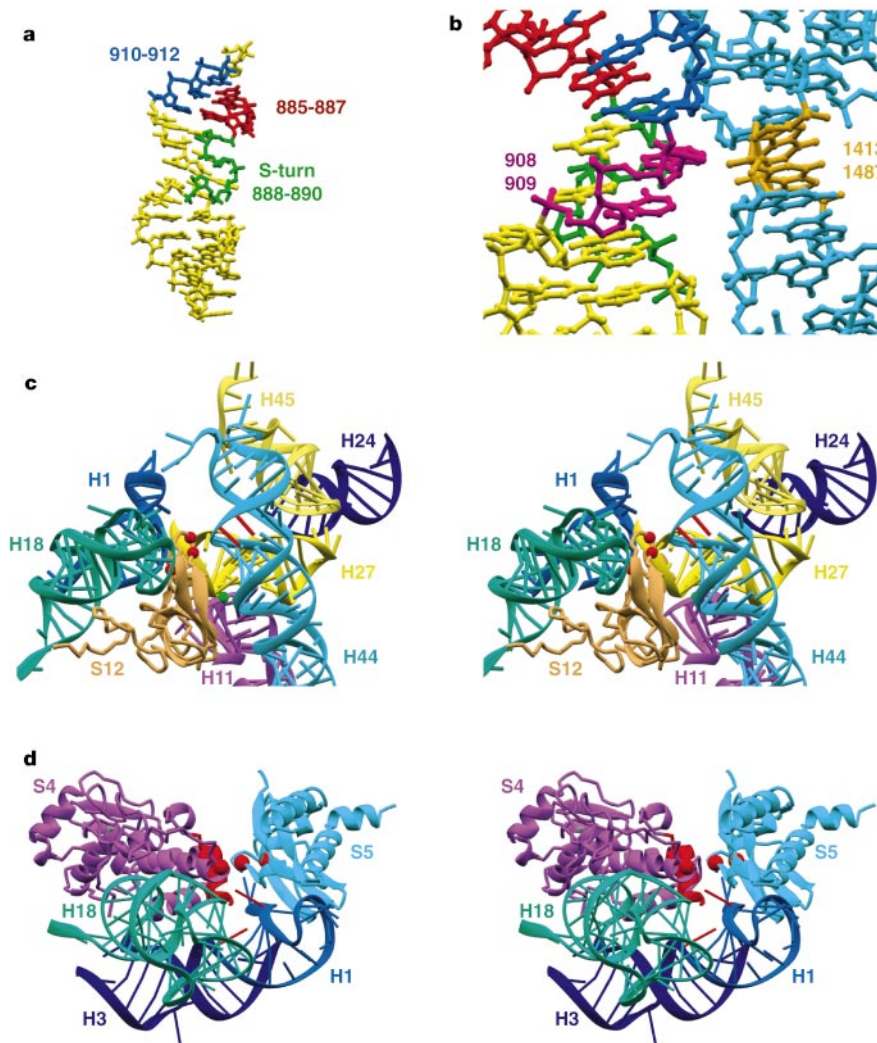


Figure 3 The H27 switch and its environment in the 30S subunit. **a**, H27 with the triplet switch highlighted in red and green. In the *ram* conformation pictured here, there is an S-turn. **b**, The packing of H27 and H44. Residues whose chemical reactivity changes in restrictive S12 strains are labelled¹⁰. **c**, The environment of H27 in the 30S subunit.

Red balls highlight sites of restrictive mutations in S12; red sticks show A1492 and A1493. **d**, The H1–H18 interaction surface is contiguous with the S4–S5 surface. The locations of *ram* mutations in S4 and S5 are highlighted in red, as are the locations of residues 8 and 26 whose chemical reactivity changes in S4 *ram* mutants¹⁰.

of H11 (Fig. 3c). Thus, it seems probable that the switch also affects other helix–helix interactions, and in fact there could be a network of coupled switches among H1, H27, H18, H34 and H44. In support of this, chemical protection experiments show that a number of RNA residues in these helices have greater chemical reactivity in the restrictive state, which indicates that the restrictive state may be a more ‘open’ state of the 30S subunit¹¹.

ram mutations often occur in S4 or S5 (reviewed in ref. 22), all but one of which map to the interface between the two proteins (red amino acids in Fig. 3d). S4 *ram* strains exhibit increased chemical reactivity in two RNA sites at the interface between the central pseudoknot (H1) and the 530 pseudoknot (H18; red RNA sticks in Fig. 1d)¹⁰. Several streptomycin resistance mutations are also found at this RNA–RNA interface. Interestingly, the S4–S5 and the H1–H18 interaction surfaces are contiguous, which indicates that this hybrid RNA–protein surface may move, and that an important function of the S4–S5 interaction may be to modulate the interaction of H1 with H18. It is probable that *ram* mutations preferentially destabilize the restrictive state, as with the S4–S5 interface and its RNA extension.

Interaction of antibiotics with the 30S

Spectinomycin. Spectinomycin inhibits elongation-factor-G-catalysed translocation of the peptidyl-tRNA from the A-site to the P-site²³. The fused ring system in spectinomycin makes it a rigid molecule. It binds in the minor groove at one end of H34, makes a single contact with a 2' OH and makes hydrogen bonds to a number of bases (Fig. 4). The most interactions are made with G1064 and C1192, consistent with protection studies²⁴ and mutagenesis data on spectinomycin resistance²⁵. These two bases are too far apart to form Watson–Crick base pairs, but are able to make a single hydrogen bond. A loop of S5 and part of H28 of 16S RNA are within 5 Å of the spectinomycin-binding site, but in this state do not make direct contacts with it. It is possible, however, that in other conformations of the 30S spectinomycin is in more direct contact with these regions.

Translocation of tRNA from one site to the next must involve movement of elements of the head (Fig. 2). It is probable that such movements involve H34 and a possible rearrangement of the connections between it and helices H35 and H38. The structure indicates that the rigid spectinomycin molecule binds near this pivot point of the head and sterically blocks movement, although it is also possible that it prevents changes in the conformation of H34. Mutations in S5 that cause resistance to spectinomycin²⁶ do not make direct contacts with the antibiotic. Rather, they map to a loop that stabilizes the interaction between the central pseudoknot, H28 and the H35–H36 region that is directly connected to H34. An attractive hypothesis is that the mutants destabilize this interaction and, by thus removing the network of interactions that stabilizes the conformation of the head to the body through S5, allows it to move even when spectinomycin is bound.

Streptomycin. Early experiments indicated that streptomycin made ribosomes error prone by affecting the proof-reading step²⁷. More recent data indicate interference with both the initial selection and proof-reading^{28–30}.

Streptomycin is tightly bound to the phosphate backbone of 16S RNA from four different parts of the molecule through both salt bridges and hydrogen bonds (Fig. 5). It also makes contact with K45 from protein S12. The four nucleotides of 16S RNA (13, 526, 915, and 1490) have been implicated in streptomycin binding on the basis of protection²⁴, crosslinking³¹ and mutagenesis data^{32–34}.

The tight interactions observed for streptomycin indicate that it preferentially stabilizes the *ram* state seen in the crystal structure. The restrictive A-site has a low tRNA affinity, whereas the *ram* state has a higher affinity^{7,11}. Therefore, by stabilizing the *ram* state, streptomycin would be expected to increase initial binding of non-cognate tRNAs. The preferential stabilization of the *ram* state would also make the transition to the restrictive state more difficult, thereby also affecting proof-reading. Thus, our results offer a structural rationale for the observed properties of streptomycin.

This stabilization of the *ram* state by streptomycin indicated by our structure can also explain much of the genetic data on the

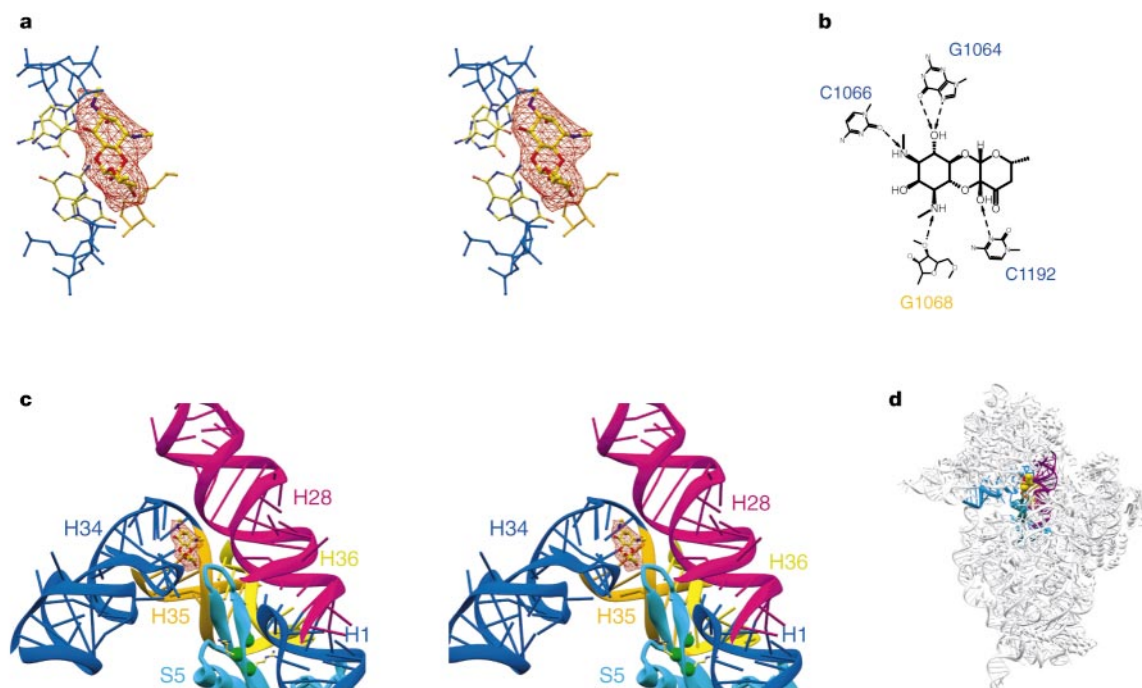


Figure 4 Interaction of spectinomycin with the 30S ribosomal subunit. **a**, Difference Fourier maps showing the binding site of spectinomycin in helix 34. **b**, Chemical structure of spectinomycin, showing interactions of the various groups with specific residues of

30S. **c**, The spectinomycin-binding site, showing its location at a pivotal point in the head of the 30S subunit. **d**, Inset showing spectinomycin in a space-filling model, and the location of its binding site on the 30S.

antibiotic. Mutations in S12 lead to a hyperaccurate phenotype (reviewed in ref. 22). A weak phenotype manifests itself as streptomycin resistance, whereas a strong phenotype (often the result of multiple mutations) leads to streptomycin dependence. Most of these mutants are to varying degrees more hyperaccurate and slower than wild-type ribosomes, consistent with destabilization of the *ram* state with respect to the restrictive state.

All the mutations in S12 map to protein loops that connect and hold in place the 908–915 and 524–527 regions, with the exception of one mutant K56 (*E. coli* K53) which contacts H44 (Fig. 2). Thus, S12 stabilizes the same region that is stabilized by streptomycin. In the resistance mutations, the *ram* state is destabilized sufficiently so that the additional stabilization induced by streptomycin does not trap the ribosome in this state. In the streptomycin-dependent mutants, the *ram* state is so destabilized that the restrictive form predominates. Streptomycin can then help stabilize the *ram* state sufficiently to restore the balance between the two states and help restore translation.

This hypothesis is supported by analysis of the K45R (*E. coli* K42) mutant, which is the only known mutant that is resistant to streptomycin but not hyperaccurate²². K45 forms a salt bridge with phosphate A913 and thus contributes to stabilization of the *ram* state. It also makes direct hydrogen-bonding contacts to two OH groups on streptomycin (Fig. 2). Mutation of this lysine to arginine would disrupt the hydrogen bonding and thereby reduce the affinity of the 30S for streptomycin, leading to resistance. However, the mutation would leave the salt bridge intact, so that the *ram* form is not destabilized and thus translation remains normal.

A number of mutations in ribosomal RNA also lead to hyperaccuracy^{32,33,35–38}. Some of these nucleotides are involved in hydrogen-bonding interactions in regions close to the streptomycin-binding site. Thus, the mutations disrupt interactions that help to stabilize the *ram* state. Others such as A915 make no contacts

with any other bases. It is possible that mutation of this base leads to more favourable contacts in the restrictive state, acting by stabilizing the restrictive state rather than destabilizing the *ram* state.

We propose that *ram* mutations in S4 and S5 preferentially destabilize the restrictive state, shifting the balance towards the *ram* state. The observation that *ram* mutations increase the affinity of ribosomes for streptomycin³⁹ is consistent with this model. Our results provide a structural basis for the notion that a delicate balance exists between the *ram* and restrictive states for optimal translation, and also explains how disruption of this balance leads to the various phenotypes observed. A definitive test of this model must await an atomic-resolution structure of the restrictive form.

Paromomycin. Paromomycin is a member of the aminoglycoside family of antibiotics, and increases the error rate of the ribosome. This family is thought to reduce the dissociation rate of A-site tRNA from the ribosome²⁸, but recent experiments also indicate that it increases the initial binding affinity of tRNA⁷.

Paromomycin binds in the major groove of H44 (Fig. 6) in a location that is in agreement with mutagenesis and protection data^{24,40}. Paromomycin ring IV contacts the backbone of both sides of H44, and ring III makes only weak contacts with the RNA. Ring II forms tight interactions with both bases and backbone of the RNA, and ring I inserts into the RNA helix and helps to flip out bases A1492 and A1493, which are not well ordered in the structure of the antibiotic-free 30S¹². Ring I mimics a nucleotide base, stacking against G1491 and hydrogen-bonding with A1408. In addition it forms a tight hydrogen-bond interaction with the phosphate backbone of A1493 which helps lock the flipped-out bases in place. The structural basis for resistance mutations has been described from the NMR structure of a complex of paromomycin with an RNA fragment corresponding to its binding site⁴¹, which is generally similar to our structure. The main differences are that ring IV is dynamically disordered in the NMR structure, and A1492 and A1493 are flipped out to a far greater extent in our structure (Fig. 6).

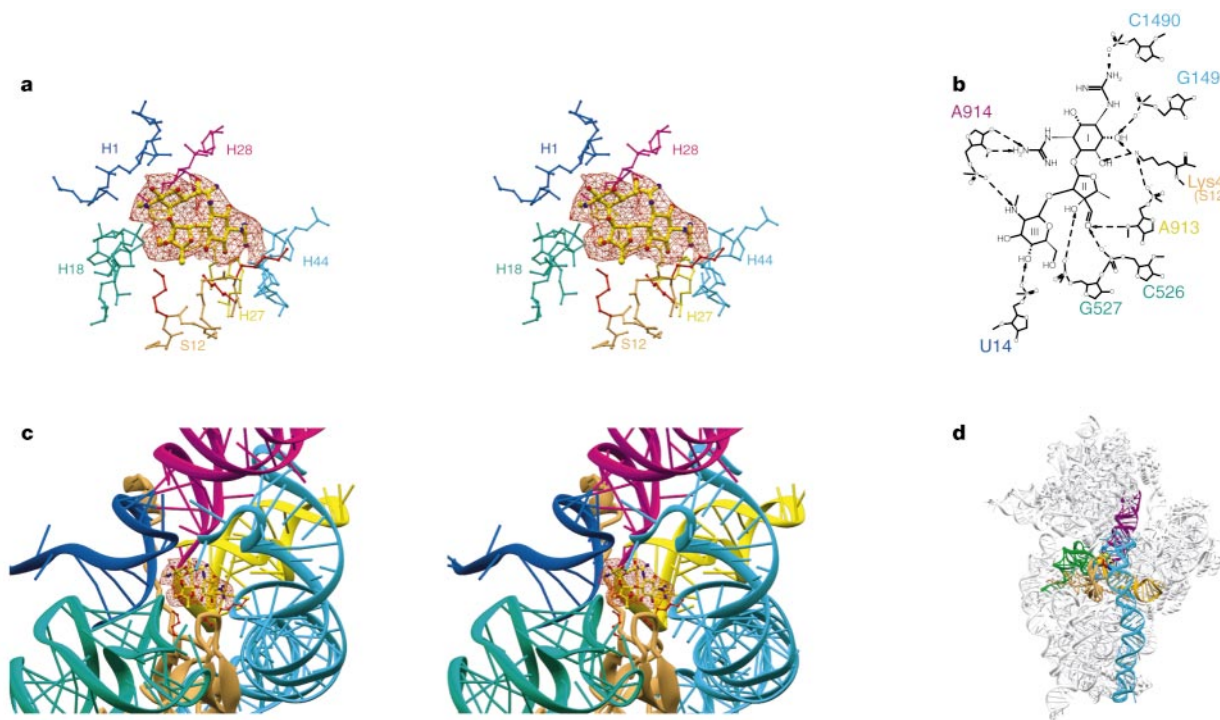


Figure 5 Interaction of streptomycin with the 30S ribosomal subunit. **a**, Difference Fourier maps showing the binding site of streptomycin. Mutations in ribosomal protein S12 that confer resistance are shown in red. **b**, Chemical structure of streptomycin, showing interactions of the various groups with specific residues of the ribosome. **c**, The

streptomycin-binding site, showing its interaction with H27, the 530 loop (H18), H44 and ribosomal protein S12. **d**, A view of the 30S showing streptomycin in a space-filling model, and the surrounding RNA and protein elements.

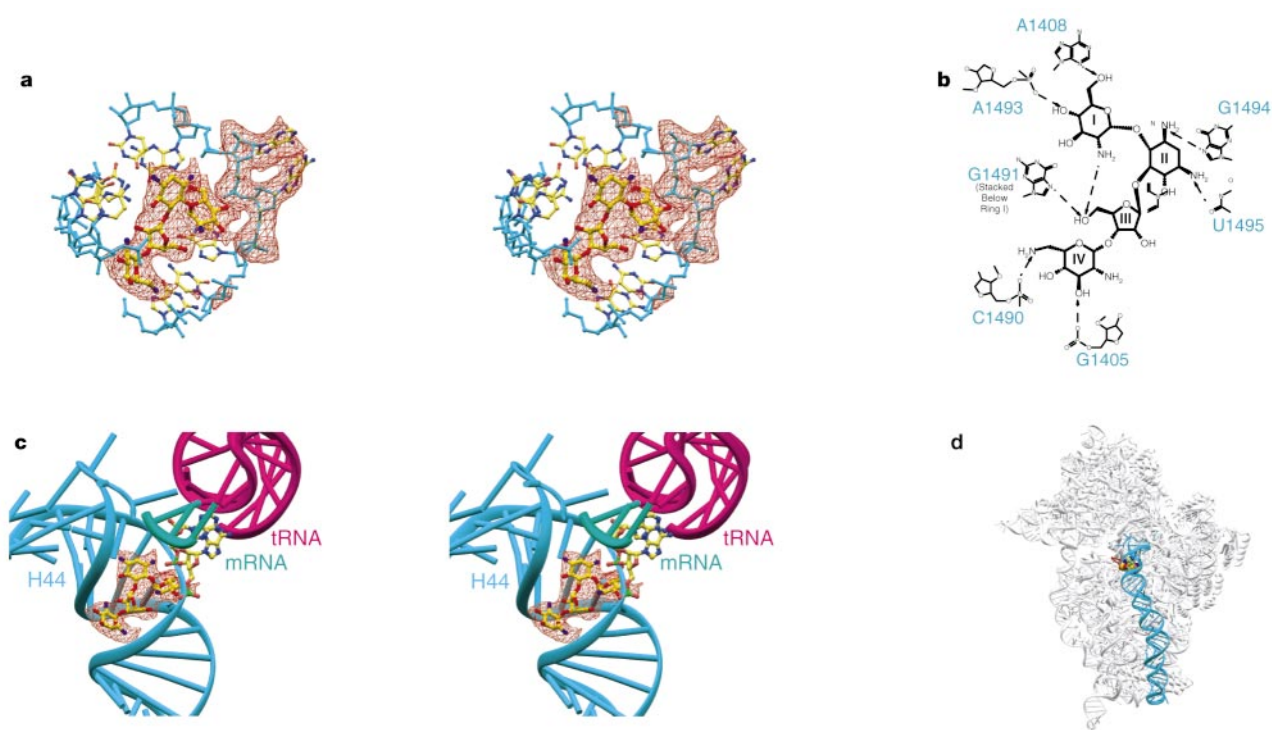


Figure 6 Interaction of paromomycin with the 30S ribosomal subunit. **a**, Difference Fourier maps showing paromomycin in a pocket formed by the major groove of H44. The bases A1492 and A1493 are flipped out from the starting model (also seen in the difference density). **b**, Chemical diagram for paromomycin, showing interactions of the various groups with specific residues of the 30S. **c**, The paromomycin-binding site, with a

model for the A-site codon and anticodon as described in Fig. 1. The bases A1492 and A1493 flipped out by paromomycin interact with the minor groove of the codon–anticodon duplex. **d**, A view of the 30S showing paromomycin in a space-filling model, and H44.

Consequently we do not see the base pair between A1408 and A1493 that was observed in the NMR structure.

Rings I and II of paromomycin are found in a number of other antibiotics including gentamycin. An NMR structure of gentamycin bound to the same fragment of H44 showed that these two rings interact with RNA in the same way as in paromomycin⁴¹. This indicates that many other aminoglycosides that bind to the decoding centre on H44 may induce errors in translation by the same mechanism as paromomycin.

With respect to the model of the A-site codon and tRNA ASL described above, the flipped-out bases point directly into the A-site and are positioned to interact with the minor groove of the codon–anticodon helix (Fig. 1b). In this modelling, there are rather strict steric constraints on the A-site anticodon and especially the codon, which is also covalently attached to both the P-site codon and downstream message. Thus, despite the uncertainties associated with modelling, it appears unlikely that the A1492–A1493 could interact with any portion of the codon–anticodon helix other than its minor groove. This model provides clues as to how paromomycin increases the affinity of the A-site for tRNA. It seems probable that in the absence of paromomycin some energy is required to flip out A1492 and A1493 so they can contact the tRNA, and presumably this energetic cost is compensated by the formation of favourable interactions with tRNA. By binding to H44, paromomycin forms a structure in which these bases are already flipped out, thus reducing the energetic cost of both cognate and non-cognate tRNA binding and increasing tRNA affinity for the A-site.

Implications for decoding in normal translation

The universally conserved residues A1492 and A1493 are required for viability in *E. coli*⁴², and are footprinted by the A-site codon and tRNA¹⁷. Experimental evidence for hydrogen bonding between the N1 atoms of A1492 and A1493 and 2' OH groups on the message⁴²

as well as modelling efforts^{42,43} have implicated these bases in the decoding process, but it has not been clear how hydrogen bonding to mRNA alone could discriminate sufficiently between cognate and non-cognate tRNAs. Moreover, in the structure of the 70S with tRNA and mRNA¹⁴, these bases were considered too distant from the codon–anticodon helix to have a direct role in decoding. A flipped-out conformation for A1492 and A1493 implies a decoding model that can reconcile these conflicting views.

The crux of decoding is the discrimination of near-cognate from cognate tRNAs. In the 30S structure, with the A-site codon and anticodon modelled as described above, the flipped-out A1492 and A1493 lie in the minor groove of the codon–anticodon helix, and form a portion of the decoding surface (Fig. 7a). We propose that the A1492 and A1493 adenine bases hydrogen-bond simultaneously to 2' OH groups on both sides of the codon–anticodon helix, in a structural motif that is seen elsewhere in the 30S structure (Fig. 7b, c). Two adenines can monitor a large portion of the minor groove of three consecutive base pairs. As the distance between the 2' OH groups across the groove is a function of the base-pairing geometry, the simultaneous formation of hydrogen bonds to both strands of the codon–anticodon helix should be sensitive to distortions arising from mispairing. We propose that the remainder of the decoding surface is composed of conserved RNA residues from the 530 pseudoknot and/or H34, which are nearby (Fig. 7a) and may monitor the shape of the correct anticodon–codon helix and contribute additional bonds. Significantly, the pairs of adenines that recognize minor grooves elsewhere in the 30S structure often recruit other RNA residues (blue in Fig. 7b and c) to complete an interaction surface.

There are at least two possible ways that adenines can hydrogen bond simultaneously to 2' OH groups on both strands of a regular helix. The docking may involve hydrogen bonding of either adenine N1, N6 and N3 (Fig. 7b), or of adenine N1, N6 and N7 (Fig. 7c). The

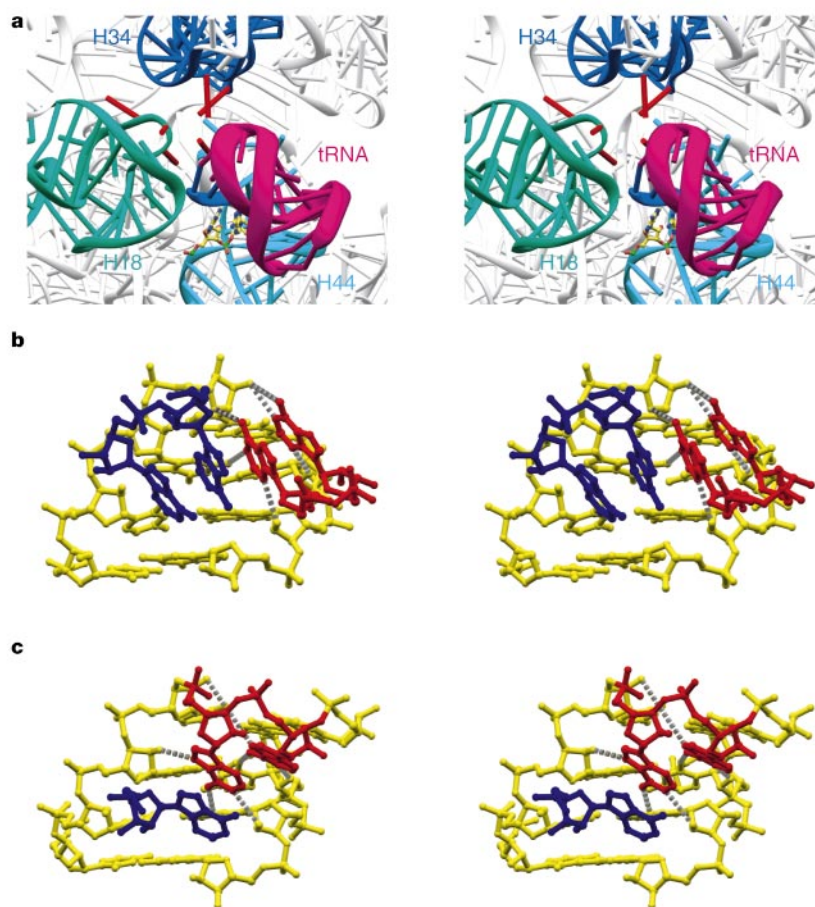


Figure 7 Model for the role of A1492 and A1493 in decoding. **a**, Overview of the 30S RNA elements surrounding the A-site helix of ASL and codon (magenta and blue, respectively), modelled as in Fig. 1. The RNA elements of the decoding site are the 530 loop (H18, green), H34 (dark blue) and the 1400–1500 region of H44 (cyan). A1492 and A1493 are represented by ball-and-stick models in the flipped-out conformation. Some of the highly conserved residues in H18 and H34 (red) that may help A1492–A1493 recognize the codon–anticodon duplex are highlighted. **b**, **c**, Two slightly different models for the

structural role of A1492–A1493 in decoding, both taken from elsewhere in the 30S structure. In both cases two adenines measure cross-strand 2'-OH – 2'-OH distances for three successive base pairs. Blue residues are other residues that help complete coverage of the minor groove. In **b**, the orientation of the adenines leaves the N7 atom pointing out and the C2 position pointing into the minor groove, and in **c** the adenine orientation leaves the N7 atom pointing in and the C2 position pointing out.

former mode would result in a steric clash if the adenine were mutated to a G and the latter would not. The distinction is important because experiments supporting a model in which A1492 and A1493 hydrogen-bond to the 2' OH of the codon made use of adenine to guanine mutants, which seem to rule out the N1–N6–N3 mode⁴². Another argument favouring the N1–N6–N7 mode is that it allows formation of a sequence-independent hydrogen bond between the adenine N6 atom and the O2 or N3 hydrogen-bond acceptor in the codon–anticodon helix⁴³ (Fig. 7c). However, the relative orientations of our flipped-out A1492–A1493 residues and the A-site codon–anticodon duplex appears to be more consistent with the N1–N6–N3 mode. The third or 'wobble' position of the codon–anticodon helix could be monitored less stringently by the A1492–A1493 in this scheme, thus allowing for the greater freedom for base pairing at the wobble position. A conclusive test of the decoding model will require the determination of the structure of cognate tRNA and mRNA ligands bound to the A-site in the restrictive form of the ribosome.

Conclusions

In this work, we have analysed the functional implications of the 30S structure reported in the accompanying paper¹². We have identified molecular details of the interactions made by A-, P- and E-site codons and tRNA with the 30S. We have also determined the structure of the 30S in complex with three antibiotics that target

different regions of the 30S ribosomal subunit and work in different ways. The detailed knowledge of their binding sites from the structure of their complex with the 30S subunit should help the design of novel drugs that target bacterial protein synthesis. It is striking that all three antibiotics bind at the functional centre, with streptomycin and paromomycin being particularly close to each other. Another common theme is that all three antibiotics work by altering a delicate balance between conformational states of the 30S as it goes through the process of translation. Each of them sheds light on a different but fundamental process during translation: spectinomycin on translocation, streptomycin on the switch between the *ram* and restrictive states, and paromomycin on decoding. Finally, the structure indicates a direct role for A1492 and A1493 in the decoding process. □

Methods

Purification of ribosomal subunits, followed by crystallization, cryoprotection and flashcooling were carried out as described¹², except that crystallization was carried out in the presence of a 10-fold molar excess of each of a mixture of streptomycin, paromomycin and spectinomycin. Also, no cobalt hexammine was included in the cryoprotectant, and antibiotics were present throughout the transfers into cryoprotectant solutions.

Data were collected on beamline ID14-4 at the ESRF in Grenoble, and integrated and scaled using HKL-2000 (ref. 44). The refined 3 Å structure of the 30S with the cobalt ions removed¹² was used as a starting model for refinement using CNS⁴⁵. Positional refinement using energy minimization was followed by grouped B-factor refinement. For cross-validation, 5% of the reflections were left out of the refinement, and care was taken to

ensure that these corresponded to the same 5% that were omitted for the refinement of the original 30S model.

Received 14 July; accepted 10 August 2000.

- Green, R. & Noller, H. F. Ribosomes and translation. *Annu. Rev. Biochem.* **66**, 679–716 (1997).
- Agrawal, R. K. *et al.* Direct visualization of A-, P-, and E-site transfer RNAs in the *Escherichia coli* ribosome. *Science* **271**, 1000–1002 (1996).
- Stark, H. *et al.* Arrangement of tRNAs in pre- and posttranslocational ribosomes revealed by electron cryomicroscopy. *Cell* **88**, 19–28 (1997).
- Kurland, C. G. Translational accuracy and the fitness of bacteria. *Annu. Rev. Genet.* **26**, 29–50 (1992).
- Pape, T., Wintermeyer, W. & Rodnina, M. V. Complete kinetic mechanism of elongation factor Tu-dependent binding of aminoacyl-tRNA to the A site of the *E. coli* ribosome. *EMBO J.* **17**, 7490–7497 (1998).
- Pape, T., Wintermeyer, W. & Rodnina, M. Induced fit in initial selection and proof-reading of aminoacyl-tRNA on the ribosome. *EMBO J.* **18**, 3800–3807 (1999).
- Pape, T., Wintermeyer, W. & Rodnina, M. V. Conformational switch in the decoding region of 16S rRNA during aminoacyl-tRNA selection on the ribosome. *Nature Struct. Biol.* **7**, 104–107 (2000).
- Gabashvili, I. S. *et al.* Major rearrangements in the 70S ribosomal 3D structure caused by a conformational switch in 16S ribosomal RNA. *EMBO J.* **18**, 6501–6507 (1999).
- Nierhaus, K. H. The allosteric three-site model for the ribosomal elongation cycle: features and future. *Biochemistry* **29**, 4997–5008 (1990).
- Allen, P. N. & Noller, H. F. Mutations in ribosomal proteins S4 and S12 influence the higher order structure of 16S ribosomal RNA. *J. Mol. Biol.* **208**, 457–468 (1989).
- Lodmell, J. S. & Dahlberg, A. E. A conformational switch in *Escherichia coli* 16S ribosomal RNA during decoding of messenger RNA. *Science* **277**, 1262–1267 (1997).
- Wimberly, B. T. *et al.* Structure of the 30S ribosomal subunit. *Nature* **407**, 327–339 (2000).
- Mueller, F. & Brimacombe, R. A new model for the three-dimensional folding of *Escherichia coli* 16S ribosomal RNA. I. Fitting the RNA to a 3D electron microscopic map at 20 Å. *J. Mol. Biol.* **271**, 524–544 (1997).
- Cate, J. H., Yusupov, M. M., Yusupova, G. Z., Earnest, T. N. & Noller, H. F. X-ray crystal structures of 70S ribosome functional complexes. *Science* **285**, 2095–2104 (1999).
- Frank, J. & Agrawal, R. K. A ratchet-like inter-subunit reorganization of the ribosome during translocation. *Nature* **406**, 319–322 (2000).
- Stark, H., Rodnina, M. V., Wieden, H. J., van Heel, M. & Wintermeyer, W. Large-scale movement of elongation factor G and extensive conformational change of the ribosome during translocation. *Cell* **100**, 301–309 (2000).
- Moazed, D. & Noller, H. F. Binding of tRNA to the ribosomal A and P sites protects two distinct sets of nucleotides in 16S rRNA. *J. Mol. Biol.* **211**, 135–145 (1990).
- Prince, J. B., Taylor, B. H., Thurlow, D. L., Ofengand, J. & Zimmermann, R. A. Covalent crosslinking of tRNA^{Val} to 16S RNA at the ribosomal P site: identification of crosslinked residues. *Proc. Natl Acad. Sci. USA* **79**, 5450–5454 (1982).
- von Ahsen, U. & Noller, H. F. Identification of bases in 16S rRNA essential for tRNA binding at the 30S ribosomal P site. *Science* **267**, 234–237 (1995).
- Mueller, F., Stark, H., van Heel, M., Rinke-Appel, J. & Brimacombe, R. A new model for the three-dimensional folding of *Escherichia coli* 16S ribosomal RNA. III. The topography of the functional centre. *J. Mol. Biol.* **271**, 566–587 (1997).
- Garrett, R. A. (eds) *The Ribosome. Structure, Function, Antibiotics and Cellular Interactions* (ASM, Washington DC, 2000).
- Kurland, C. G., Hughes, D., Ehrenberg, M. in *Limitations of Translational Accuracy* (eds Neidhardt, F. C. *et al.*) 979–1003 (ASM, Washington DC, 1996).
- Bilgin, N., Richter, A. A., Ehrenberg, M., Dahlberg, A. E. & Kurland, C. G. Ribosomal RNA and protein mutants resistant to spectinomycin. *EMBO J.* **9**, 735–739 (1990).
- Moazed, D. & Noller, H. F. Interaction of antibiotics with functional sites in 16S ribosomal RNA. *Nature* **327**, 389–394 (1987).
- Brink, M. F., Brink, G., Verbeet, M. P. & de Boer, H. A. Spectinomycin interacts specifically with the residues G1064 and C1192 in 16S rRNA, thereby potentially freezing this molecule into an inactive conformation. *Nucleic Acids Res.* **22**, 325–331 (1994).
- Wittmann-Liebold, B. & Greuer, B. The primary structure of protein S5 from the small subunit of the

Escherichia coli ribosome. *FEBS Lett.* **95**, 91–98 (1978).

- Ruusala, T. & Kurland, C. G. Streptomycin preferentially perturbs ribosomal proofreading. *Mol. Gen. Genet.* **198**, 100–104 (1984).
- Karimi, R. & Ehrenberg, M. Dissociation rate of cognate peptidyl-tRNA from the A-site of hyper-accurate and error-prone ribosomes. *Eur. J. Biochem.* **226**, 355–360 (1994).
- Karimi, R. & Ehrenberg, M. Dissociation rates of peptidyl-tRNA from the P-site of *E. coli* ribosomes. *EMBO J.* **15**, 1149–1154 (1996).
- Bilgin, N. & Ehrenberg, M. Mutations in 23S ribosomal RNA perturb transfer RNA selection and can lead to streptomycin dependence. *J. Mol. Biol.* **235**, 813–824 (1994).
- Gravel, M., Melancon, P. & Brakier-Gingras, L. Cross-linking of streptomycin to the 16S ribosomal RNA of *Escherichia coli*. *Biochemistry* **26**, 6227–6232 (1987).
- Montandon, P. E., Wagner, R. & Stutz, E. *E. coli* ribosomes with a C912 to U base change in the 16S rRNA are streptomycin resistant. *EMBO J.* **5**, 3705–3708 (1986).
- Pinar, R., Payant, C., Melancon, P. & Brakier-Gingras, L. The 5' proximal helix of 16S rRNA is involved in the binding of streptomycin to the ribosome. *FASEB J.* **7**, 173–176 (1993).
- Melancon, P., Lemieux, C. & Brakier-Gingras, L. A mutation in the 530 loop of *Escherichia coli* 16S ribosomal RNA causes resistance to streptomycin. *Nucleic Acids Res.* **16**, 9631–9639 (1988).
- Montandon, P. E., Nicolas, P., Schurmann, P. & Stutz, E. Streptomycin-resistance of *Euglena gracilis* chloroplasts: identification of a point mutation in the 16S rRNA gene in an invariant position. *Nucleic Acids Res.* **13**, 4299–4310 (1985).
- Leclerc, D., Melancon, P. & Brakier-Gingras, L. Mutations in the 915 region of *Escherichia coli* 16S ribosomal RNA reduce the binding of streptomycin to the ribosome. *Nucleic Acids Res.* **19**, 3973–3977 (1991).
- Melancon, P., Boileau, G. & Brakier-Gingras, L. Cross-linking of streptomycin to the 30S subunit of *Escherichia coli* with phenyldiglyoxal. *Biochemistry* **23**, 6697–6703 (1984).
- Powers, T. & Noller, H. F. A functional pseudoknot in 16S ribosomal RNA. *EMBO J.* **10**, 2203–2214 (1991).
- Bock, A., Petzet, A. & Piepersberg, W. Ribosomal ambiguity (ram) mutations facilitate dihydrostreptomycin binding to ribosomes. *FEBS Lett.* **104**, 317–321 (1979).
- Spahn, C. M. & Prescott, C. D. Throwing a spanner in the works: antibiotics and the translation apparatus. *J. Mol. Med.* **74**, 423–439 (1996).
- Fourmy, D., Recht, M. I., Blanchard, S. C. & Puglisi, J. D. Structure of the A site of *Escherichia coli* 16S ribosomal RNA complexed with an aminoglycoside antibiotic. *Science* **274**, 1367–1371 (1996).
- Yoshizawa, S., Fourmy, D. & Puglisi, J. D. Recognition of the codon–anticodon helix by ribosomal RNA. *Science* **285**, 1722–1725 (1999).
- VanLoock, M. S., Easterwood, T. R. & Harvey, S. C. Major groove binding of the tRNA/mRNA complex to the 16S ribosomal RNA decoding site. *J. Mol. Biol.* **285**, 2069–2078 (1999).
- Otwinowski, Z. & Minor, W. in *Methods in Enzymology* (eds Carter, C. W. J. & Sweet, R. M.) 307–325 (Academic, New York, 1997).
- Brünger, A. T. *et al.* Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr D* **54**, 905–921 (1998).
- Carson, M. Ribbons 2.0. *J. Appl. Cryst.* **24**, 958–961 (1991).

Supplementary information is available on *Nature's* World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of *Nature*.

Acknowledgements

This work was supported by the Medical Research Council (UK) and a grant from the US National Institutes of Health to V.R. and S.W. White, D.E.B. was supported by an EMBO long-term fellowship, and W.M.C. by an NIH predoctoral fellowship. We thank R. Ravelli for help with data collection, and M. Pacold for help and discussions with the modelling of antibiotic structures.

Correspondence and requests for materials should be addressed to V.R. (e-mail: ramak@mrc-lmb.cam.ac.uk) or B.T.W. (e-mail: brianw@mrc-lmb.cam.ac.uk). Coordinates for the antibiotic–30S complex have been deposited in the Protein Data Bank, accession code 1FJG.

The accelerations of stars orbiting the Milky Way's central black hole

A. M. Ghez, M. Morris, E. E. Becklin, A. Tanner & T. Kremenek

Department of Physics and Astronomy, UCLA, Los Angeles, California 90095-1562, USA

Recent measurements^{1–4} of the velocities of stars near the centre of the Milky Way have provided the strongest evidence for the presence of a supermassive black hole in a galaxy⁵, but the observational uncertainties poorly constrain many of the black hole's properties. Determining the accelerations of stars in their orbits around the centre provides much more precise information about the position and mass of the black hole. Here we report measurements of the accelerations of three stars located ~ 0.005 pc (projected on the sky) from the central radio source Sagittarius A* (Sgr A*); these accelerations are comparable to those experienced by the Earth as it orbits the Sun. These data increase the inferred minimum mass density in the central region of the Galaxy by an order of magnitude relative to previous results, and localize the dark mass to within 0.05 ± 0.04 arcsec of the nominal position of Sgr A*. In addition, the orbital period of one of the observed stars could be as short as 15 years, allowing us the opportunity in the near future to observe an entire period.

In 1995, we initiated a programme of high-resolution $2.2\text{-}\mu\text{m}$ (K band) imaging of the inner $5\text{ arcsec} \times 5\text{ arcsec}$ of the Galaxy's central stellar cluster with the W. M. Keck 10-m telescope on Mauna Kea, Hawaii ($1\text{ arcsec} = 0.04\text{ pc}$ at the distance to the Galactic Centre, 8 kpc , ref. 6). From each observation, several thousand short exposure ($t_{\text{exp}} = 0.137\text{ s}$) frames were collected, using the facility near-infrared camera, NIRC^{7,8}, and combined to produce a final diffraction-limited image having an angular resolution of 0.05 arcsec . Between 1995 and 1997, images were obtained once a year with the aim of detecting the stars' velocities in the plane of the sky. The results from these measurements are detailed in ref. 4; in summary, two-dimensional velocities were measured for 90 stars with simple linear fits to the positions as a function of time. These velocities, which reach up to $1,400\text{ km s}^{-1}$, implied the existence of a $2.6 \times 10^6 M_{\odot}$ black hole coincident ($\pm 0.1\text{ arcsec}$) with the nominal location of Sgr A* (ref. 9), the unusual radio source^{10–12} long believed to be the counterpart of the putative black hole.

The new observations presented here were obtained several times a year from 1997 to 1999 to improve the sensitivity to accelerations. With nine independent measurements, we now fit the positions of stars as a function of time with second order polynomials (see Fig. 1). The resulting velocity uncertainties are reduced by a factor of 3 compared to our earlier work, primarily as a result of the increased time baseline, and, in the central square arcsecond, by a factor of 6 compared to that presented in ref. 3, owing to our higher angular resolution. Among the 90 stars in our original proper motion

sample⁴, we have now detected significant accelerations for three stars, SO-1, SO-2 and SO-4 (see Table 1); specifically, these are the sources for which the reduced chi-squared value for a quadratic fit is smaller than that for the linear model of their motions by more than 1. These three stars are independently distinguished in our sample, being among the fastest moving stars ($v = 570$ to $1,350\text{ km s}^{-1}$) and among the closest to the nominal position of Sgr A* ($r_{1995} = 0.004\text{--}0.013\text{ pc}$). With accelerations of $2\text{--}5\text{ milliarcsec yr}^{-2}$, or equivalently $(3\text{--}6) \times 10^{-6}\text{ km s}^{-2}$, they are experiencing accelerations similar to the Earth in its orbit about the Sun.

Acceleration vectors, in principle, are more precise tools than the velocity vectors for studying the central mass distribution. Even projected onto the plane of the sky, each acceleration vector should be oriented in the direction of the central mass, assuming a spherically symmetric potential. Thus, the intersection of multiple acceleration vectors is the location of the dark mass. Figure 2 shows the acceleration vector's direction for the three stars. Within 1σ ,

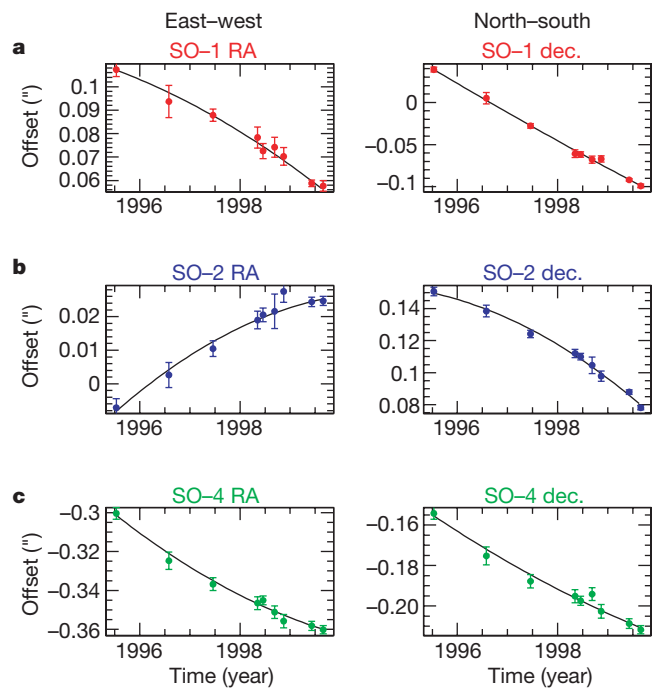


Figure 1 East–west and north–south positional offsets from the nominal location of Sgr A* versus time for (a) SO-1, (b) SO-2, and (c) SO-4. The offset range shown is scaled to the points in each plot and therefore varies from $\sim 0.4''$ to $\sim 0.15''$. Each of these stars is located with a precision of about $1\text{--}5\text{ mas}$ in the individual maps and the alignment of these positions, which is carried out by minimizing the net displacement of all stars as described in ref. 4, has an uncertainty of about 3 mas based on the half sample bootstrap resampling method²² (the alignment uncertainty is at a minimum at the centre of the map and grows linearly with radius). These two uncertainty terms are added in quadrature; the results are used in the fitting process and depicted as errorbars here. In each plot, the solid line shows the second order polynomial used to derive the acceleration term. These plots demonstrate that we are able to measure, for the first time, accelerations of 2 to 5 mas yr^{-2} ($0.3\text{--}0.6\text{ cm s}^{-2}$) for stars orbiting a supermassive black hole.

Table 1 Measurements for stars with significant accelerations

	SO-1 (S1)	SO-2 (S2)	SO-4 (S8)
Radius from Sgr A*–magnitude (milliparsecs)	4.42 ± 0.05	5.83 ± 0.04	13.15 ± 0.04
Radius from Sgr A*–position angle (degrees)	290.1 ± 0.7	3.1 ± 0.4	117.3 ± 0.2
Velocity–magnitude (km s^{-1})	1350 ± 40	570 ± 20	990 ± 30
Velocity–position angle (degrees)	168 ± 2	241 ± 2	129 ± 2
Acceleration–magnitude ($\text{milliarcsec yr}^{-2}$)	2.4 ± 0.7	5.4 ± 0.3	3.2 ± 0.5
Acceleration–position angle (degrees)	80 ± 15	154 ± 4	294 ± 7

The primary nomenclature adopted here and in the text is from ref. 4; in parentheses star names from ref. 1 are also given. All quantities listed are derived from the second-order polynomial fit to the data and are given for epoch 1995.53. All position angles are measured east of north. All uncertainties are 1σ and are determined by the jackknife resampling method²². The radius uncertainties listed include only the relative positional uncertainties and not the uncertainty in the origin used (the nominal position of Sgr A*; see text for offset to measured dynamical centre).

these vectors do indeed overlap, and furthermore, the intersection point lies a mere 0.05 ± 0.03 arcsec east and 0.02 ± 0.03 arcsec south of the nominal position of Sgr A*, consistent with the identification of Sgr A* as the carrier of the mass. Previously, with statistical treatments of velocities only, the dynamical centre was located to within ± 0.1 arcsec (1σ) of Sgr A*'s position⁴. This velocity-based measurement is unaffected by the increased time baseline, as its uncertainty is dominated by the limited number of stars at a given radius. The accelerations thus improve the localization of our Galaxy's dynamical centre by a factor of 3, which is essential for reliably associating any near-infrared source with the black hole given the complexity of the region.

Like the directions, the magnitudes of the acceleration vectors also constrain the central black hole's properties. In three dimensions, the acceleration and radius vectors (a_{3D} and r_{3D} , respectively) provide a direct measure of the enclosed mass simply by $a_{3D} = G \times M/r_{3D}^2$. For a central potential, the acceleration and radius vectors are co-aligned and, with a projection angle to the plane of the sky θ , the two-dimensional projections place the following lower limit on the central mass: $M \cos^3(\theta) = a_{2D} \times r_{2D}^2/G$. This analysis is independent of the star's orbital parameters, although θ is in fact a lower limit for the orbital inclination angle. Figure 3 shows the minimum mass implied by each star's two-dimensional acceleration as a function of projected radius, with dashed curves displaying how the implied mass and radius grow with projection angle. For each point, the uncertainty in the position of the dynamical centre dominates the minimum mass uncertainties. Also plotted are the results from the statistical analysis of the velocity vectors measured in the plane of the sky, which imply a dark mass of $(2.3\text{--}3.3) \times 10^6 M_\odot$ inside a radius of $0.015 \text{ pc}^{3,4}$. If, as has been assumed, this dark mass is in the form of a single supermassive black hole, the enclosed mass should remain level at smaller radii. Projection angles of $51\text{--}56^\circ$ and $25\text{--}37^\circ$ for S0-1 and S0-2, respectively, would yield the mass inferred from velocities and place these stars at a mere $\sim 0.008 \text{ pc}$ (solid line portions of the limiting mass curves in Fig. 3), thus increasing the dark mass density implied by velocities by an order of magnitude to $8 \times 10^{12} M_\odot \text{ pc}^{-3}$.

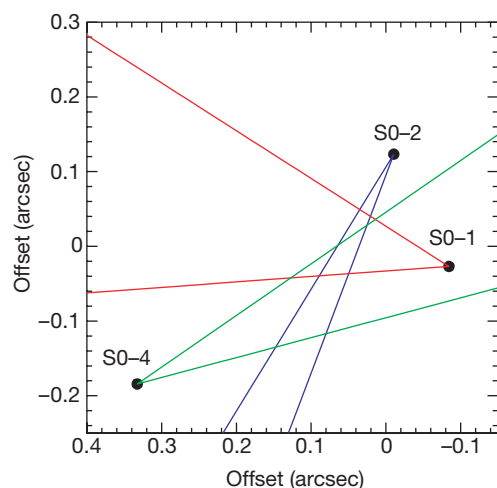


Figure 2 The acceleration uncertainty cones and their intersections. The cones' edges represent the directions for which the accelerations deviate by 1σ from their best-fit values and their vertices are the time-averaged positions, measured relative to the nominal position of Sgr A*, rather than the positions listed in Table 1. If the accelerations are caused by a single supermassive black hole, or even a spherically symmetric mass distribution, these vectors should intersect at a common location, the centre of the mass. The 1σ intersection point lies 0.05 ± 0.03 arcsec east and 0.02 ± 0.03 arcsec south of the nominal position of Sgr A*. The existence of an intersection point suggests that there is indeed a common origin for the acceleration and pinpoints the location of the black hole to within 0.03 arcsec.

With smaller projection angles, these two stars also allow for the enclosed mass to decrease at smaller radii, as would occur in the presence of an extended distribution of dark matter surrounding a less massive black hole^{13–17}. In contrast to the agreement between the mass distribution inferred from the velocities and the accelerations for S0-1 and S0-2, the minimum mass implied by S0-4's acceleration is inconsistent at least at the 1σ level. We note, in support of the validity of S0-4's acceleration vector, that its orientation is consistent with the intersection of the S0-1 and S0-2 acceleration vectors (see Fig. 2). Nonetheless, continued monitoring of S0-4 will be important for assessing this possible discrepancy. Overall, the individual magnitudes of the three acceleration vectors support the existence of a central black hole with a mass of approximately $3 \times 10^6 M_\odot$.

Using acceleration measurements, it is now possible to constrain the individual orbits. We checked the conclusions of the previous

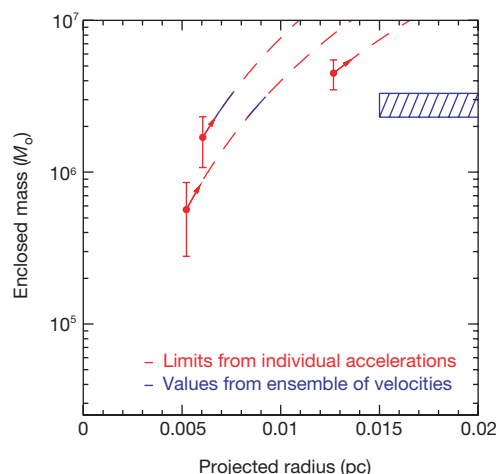


Figure 3 The minimum enclosed mass implied by each star's acceleration measurement versus projected distance from the newly determined dynamical centre position. If the projection between the plane of the sky and both the radius and the acceleration vectors is θ , then the true mass increases as $1/\cos^3(\theta)$ and the true acceleration and radius vectors increase as $1/\cos(\theta)$. Also plotted is the mass range implied from a statistical analysis of the approximately 100 velocity vectors that have been measured in the plane of the sky^{3,4}.

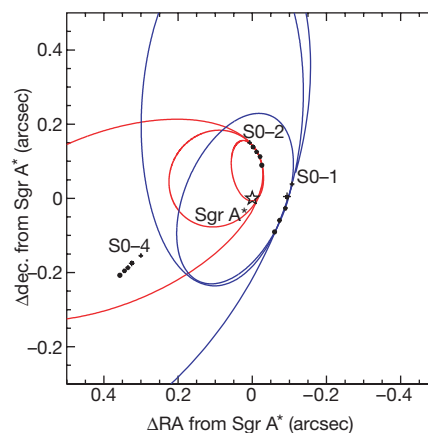


Figure 4 The measured motion of S0-1, S0-2 and S0-4 and several allowed orbital solutions. Only the measurements obtained in June of each of the 5 years are shown. S0-1 and S0-2 are moving clockwise about Sgr A* and S0-4 is travelling radially outward. In the orbital analysis, two constraints are applied, a central mass of $2.6 \times 10^6 M_\odot$ and a true focus located at the position of Sgr A*. Displayed are orbital solutions with periods of 17, 80 and 505 years for S0-2 and 63, 200 and 966 years for S0-1.

studies^{3,4} by assuming that the stars are bound to a central mass of $(2.2\text{--}3.3) \times 10^6 M_\odot$ located within 0.03 arcsec of the nominal position of Sgr A*. Excellent orbital fits were found for S0-1 and S0-2 for the entire range of masses and for true foci within 0.01 arcsec of the nominal position of Sgr A*, suggesting that we now have comparable accuracy in determining the infrared location of Sgr A* (ref. 9) and the dynamical centre of our Galaxy (see Fig. 4). The orbital solutions for these stars have eccentricities ranging from 0 (circular) to 0.9 for S0-1 and 0.5 to 0.9 for S0-2 and periods in the range 35–1,200 and 15–550 years, respectively, raising the possibility of seeing a star make a complete journey around the centre of the Galaxy within the foreseeable future. Although the fits are not yet unique, they impose a maximum orbital distance from the black hole, or apoapse, of 0.1 pc. This suggests that S0-1 and S0-2 might have formed locally. If these stars are indeed young², then their small apoapse distance presents a challenge to classical star formation theories in light of the strong tidal forces created by the central black hole and might require a collisional or compressional star formation scenario^{18,19}. However, dynamical friction may be able to act on a short enough timescale to bring these stars in from a much larger distance²⁰. More accurate orbital parameters are needed to fully address these problems and others such as the distance to the Galactic Centre²¹. The determinations of these parameters will be considerably improved when radial velocities are measured for these stars using adaptive optics techniques, as the current solutions based on the proper motion data alone predict radial velocities ranging from 200 and 2,000 km s⁻¹. □

Received 24 May; accepted 19 July 2000.

- Eckart, A. & Genzel, R. Stellar proper motions in the central 0.1 pc of the Galaxy. *Mon. Not. R. Astron. Soc.* **284**, 576–598 (1997).
- Genzel, R., Eckart, A., Ott, T. & Eisenhauer, F. On the nature of the dark mass in the centre of the Milky Way. *Mon. Not. R. Astron. Soc.* **291**, 234–249 (1997).
- Genzel, R., Pichon, C., Eckart, A., Gerhard, O. E. & Ott, T. Stellar dynamics in the galactic centre: Proper motions and anisotropy. *Mon. Not. R. Astron. Soc.* (in the press).
- Ghez, A. M., Klein, B. L., Morris, M. & Becklin, E. E. High proper-motion stars in the vicinity of Sagittarius A*: Evidence for a supermassive black hole at the center of our galaxy. *Astrophys. J.* **509**, 678–686 (1998).
- Maoz, E. Dynamical constraints on alternatives to supermassive black holes in galactic nuclei. *Astrophys. J.* **494**, L181–L184 (1998).
- Reid, M. J. The distance to the center of the Galaxy. *Annu. Rev. Astron. Astrophys.* **31**, 345–372 (1993).
- Matthews, K. & Soifer, B. T. in *Astronomy with Infrared Arrays: The Next Generation* (ed. McLean, I.) Vol. 190, 239–246 (Astrophysics and Space Sciences Library, Kluwer Academic, Dordrecht, 1994).
- Matthews, K., Ghez, A. M., Weinberger, A. J. & Neugebauer, G. The diffraction-limited images from the W. M. Keck Telescope. *Proc. Astron. Soc. Pacif.* **108**, 615–619 (1996).
- Menten, K. M., Reid, M. J., Eckart, A. & Genzel, R. The position of Sgr A*: Accurate alignment of the radio and infrared reference frames at the galactic center. *Astrophys. J.* **475**, L111–L114 (1997).
- Lo, K. Y., Shen, Z.-Q., Zhao, J.-H. & Ho, P. T. Intrinsic size of SGR A*: 72 Schwarzschild radii. *Astrophys. J.* **508**, L61–L64 (1998).
- Reid, M. J., Readhead, A. C. S., Vermeulen, R. C. & Treuhaft, R. N. The proper motion of Sagittarius A*. I. First VLBA results. *Astrophys. J.* **524**, 816–823 (1999).
- Backer, D. C. & Sramek, R. A. Proper motion of the compact, nonthermal radio source in the galactic center, Sagittarius A*. *Astrophys. J.* **524**, 805–815 (1999).
- Salati, P. & Silk, J. A stellar probe of dark matter annihilation in galactic nuclei. *Astrophys. J.* **338**, 24–31 (1989).
- Tsiklauri, D. & Viollier, R. D. Dark matter concentration in the galactic center. *Astrophys. J.* **500**, 591–595 (1998).
- Gondolo, P. & Silk, J. Dark matter annihilation at the galactic center. *Phys. Rev. Lett.* **83**, 1719–1722 (1999).
- Torres, D. F., Capozziello, S. & Lambiase, G. A supermassive boson star at the galactic center? *Astrophys. J.* (in the press).
- Romanowsky, A. & Kochanek, C. in *Proceedings of Dynamics of Star Clusters and the Milky Way* (ASP Conference Series, Astronomical Society of the Pacific, San Francisco, in the press).
- Sanders, R. H. The case against a massive black hole at the Galactic Centre. *Nature* **359**, 131–132 (1992).
- Morris, M., Ghez, A. M. & Becklin, E. E. The galactic center black hole: clues for the evolution of black holes in galactic nuclei. *Adv. Space Res.* **23**, 959–968 (1999).
- Gerhard, O. The galactic center He I stars: remains of a dissolved young cluster? *Astrophys. J.* (submitted).
- Salim, S. & Gould, A. Sagittarius A* “Visual Binaries”: A direct measurement of the galactocentric distance. *Astrophys. J.* **523**, 633–641 (1999).
- Babu, G. J. & Feigelson, E. D. *Astrostistics* (Chapman & Hall, London, 1996).

Acknowledgements

This work was supported by the National Science Foundation and the Packard Foundation. We are grateful to J. Larkin for exchanging telescope time; O. Gerhard, M. Jura, and

A. Weinberger for useful input; telescope observing assistants J. Aycok, T. Chelminiak, G. Puniwai, C. Sorenson, W. Wack, M. Whittle and software/instrument specialists A. Conrad and B. Goodrich for their help during the observations. The data presented here were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California and the National Aeronautics and Space Administration. The Observatory was made possible by the financial support of the W. M. Keck Foundation.

Correspondence and requests for materials should be addressed to A.G.
(e-mail: ghez@astro.ucla.edu).

Onset of antiferromagnetism in heavy-fermion metals

A. Schröder*, G. Aeppli†, R. Coldea‡§, M. Adams§, O. Stockert*||, H.v. Löhneysen*, E. Bucher¶#, R. Ramazashvili* & P. Coleman**

* Physikalisches Institut, Universität Karlsruhe, D-76128 Karlsruhe, Germany

† NEC, 4 Independence Way, Princeton, New Jersey 08540, USA

‡ Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831, USA

§ ISIS Facility, CCLRC, Rutherford-Appleton Laboratory, Didcot OX11 0QX, UK

¶ Department of Physics, University of Konstanz, D-78457 Konstanz, Germany

Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA

* Department of Physics, University of Illinois, Urbana, Illinois 61801, USA

** Center for Materials Theory, Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854, USA

There are two main theoretical descriptions of antiferromagnets. The first arises from atomic physics, which predicts that atoms with unpaired electrons develop magnetic moments. In a solid, the coupling between moments on nearby ions then yields antiferromagnetic order at low temperatures¹. The second description, based on the physics of electron fluids or ‘Fermi liquids’, states that Coulomb interactions can drive the fluid to adopt a more stable configuration by developing a spin density wave^{2,3}. It is at present unknown which view is appropriate at a ‘quantum critical point’, where the antiferromagnetic transition temperature vanishes^{4–7}. Here we report neutron scattering and bulk magnetometry measurements of the metal CeCu_{6-x}Au_x, which allow us to discriminate between the two models. We find evidence for an atomically local contribution to the magnetic correlations which develops at the critical gold concentration ($x_c = 0.1$), corresponding to a magnetic ordering temperature of zero. This contribution implies that a Fermi-liquid-destroying spin-localizing transition, unanticipated from the spin density wave description, coincides with the antiferromagnetic quantum critical point.

CeCu₆ (ref. 8, 9) is a ‘heavy-fermion’ compound, a class of metal formed by mixing an actinide or rare earth with a lighter element. ‘Heavy’ refers to the extremely large effective masses of the charge carriers at low temperatures, often hundreds or even thousands of times greater than the electron mass. Because of their large effective electron masses and proximity to antiferromagnetism^{10,11} in the phase diagram, heavy-fermion materials provide an opportunity to observe the competition between metallic electron band formation and magnetism. The large masses imply small bandwidths, and the magnetism is easily ‘tuned’ by external pressure¹², magnetic field and chemical composition. Thus all phenomena associated with this competition can be readily investigated at low temperatures in these materials. It is much easier to collect clean data at such

|| Present address: H.H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK.

temperatures than at the higher temperatures needed for other materials—such as the transition metals or their oxides^{7,13}, which also display competition between band formation and magnetism, but which have much higher characteristic energies. Because of a substantial existing amount of data^{14–18} and the ability to grow large and highly homogeneous single crystals, we have chosen $\text{CeCu}_{6-x}\text{Au}_x$ as a test material. The quantum critical point (QCP) occurs as gold is substituted for the Cu atoms; when more than 0.1 Cu sites per cerium are replaced¹⁹, the heavy-fermion paramagnet gives way to an ordered antiferromagnet^{20,21} (Fig. 1a).

Figure 1 illustrates two competing models of magnetism that have been widely discussed in the context of heavy-fermion systems^{22–25}. The fundamental parameter is the strength of the hybridization W between the localized f orbitals and the more extended s , p and d orbitals. When W is small, the extended orbitals form a metal which is weakly coupled to the unfilled, and hence magnetic, f orbitals. Thus, as the temperature is reduced, the two subsystems undergo

largely independent evolutions towards Fermi liquid and magnetically ordered ground states, a situation found in many elemental rare earths (such as gadolinium), and their compounds. The volume bounded by the Fermi surface is relatively small, containing only electrons from the extended orbitals. As we begin to increase W , the magnetic coupling between the rare-earth ions via the conduction electrons and the magnetic ordering (or Néel) temperature, T_N , both increase. On the other hand, if W is large, the material develops a Fermi liquid of strongly hybridized carriers with a relatively large Fermi volume containing electrons from both extended and f orbitals. The kinetic energy of band electrons then overwhelms interaction effects, with the result that T_N is suppressed to zero for W beyond a critical value W_c (ref. 10).

An important scale for heavy-fermion materials is the Kondo temperature, T_K ; the temperature below which the magnetic susceptibility saturates, and the hybridization of the f orbitals with the other orbitals becomes apparent. In contrast to T_N , the Kondo temperature increases monotonically with W . The central open question concerns how this scale behaves near the QCP: in particular, whether T_K remains finite when T_N vanishes (Fig. 1b) or whether T_K and T_N vanish at the same point (Fig. 1c). The first case occurs if the local moments are quenched at a finite temperature above the QCP: here, local moments do not play a role in the physics of the QCP and a spin-density wave model must be used^{5,23,24}. In the second case, local moments are present at all temperatures down to $T = 0$ at the QCP, so the heavy-electron Fermi surface ceases to exist at the QCP. Here we report comprehensive evidence in favour of the second case; see Fig. 1c.

To discriminate between the local-moment and spin-density-wave cases, we have measured the magnetic response function $\chi(\mathbf{q}, E) = \chi'(\mathbf{q}, E) + i\chi''(\mathbf{q}, E)$, as a function of temperature (T) and external magnetic fields (H), using neutron scattering and bulk magnetometry. The function $\chi(\mathbf{q}, E)$ describes the Fourier components at frequencies $E/\hbar$ of the magnetization $\delta m(\mathbf{q}, t) = \chi(\mathbf{q}, t)\delta h_0$ induced by a unit external field pulse $\delta h(\mathbf{r}, t) = \delta h_0 \exp(i\mathbf{q} \cdot \mathbf{r})\delta(t)$ with spatial modulation described by the wavevector $\mathbf{q}$. A simple

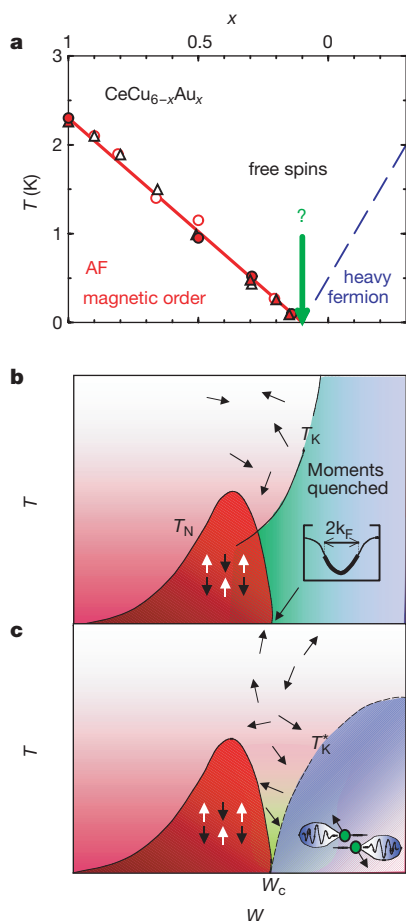


Figure 1 Phase diagram and models for quantum criticality in $\text{CeCu}_{6-x}\text{Au}_x$. **a**, Phase diagram for $\text{CeCu}_{6-x}\text{Au}_x$, where increasing Au concentration x drives the heavy-fermion alloy into an antiferromagnetically ordered (AF) state. The quantum critical point (QCP) at $x = 0.1$ is the subject of this Letter. **b, c**, Two competing models for the magnetic QCP in heavy-fermion metals. W is the strength of the hybridization between the localized f electrons and the surrounding conduction sea, which causes the antiferromagnetic order (with transition temperature T_N) to collapse at a critical value $W = W_c$. **b**, The local moments are magnetically quenched—absorbed into the metallic bands—at a finite Kondo temperature, T_K , and do not play a role at the QCP. Antiferromagnetism develops via a ‘spin-density-wave instability’ of the underlying Fermi surface. **c**, Local moments exist down to an effective lattice Kondo temperature, T_K , which vanishes at the QCP. In this model, local moments become quenched precisely at the QCP, and play an active role in the magnetic critical fluctuations.

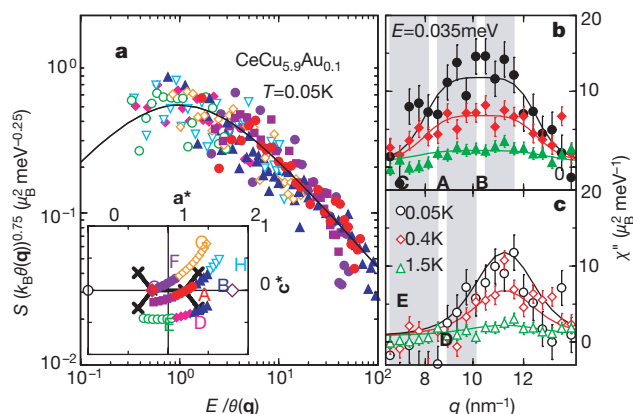


Figure 2 Magnetic neutron scattering data. **a**, Spectra at wavevectors $\mathbf{q}$ indicated in inset, as a function of energy transfer E in units of the $\mathbf{q}$ -dependent Weiss temperature $\theta(\mathbf{q})$. Intensities are correspondingly normalized to $\theta(\mathbf{q})^{-\alpha}$ ($\alpha = 0.75$), as suggested by equation (4). The line is the scaling function $G(\delta) = C/(1 + (-i\delta/a)^\alpha)$ with $\delta = E/\theta(\mathbf{q})$ and $a = 0.8k_B$. The inset is a map of reciprocal space showing the critical line (black lines) close to the magnetic Bragg peaks in the ordered regime (black crosses) and the wavevector regions that we consider (labelled A–H). **b, c**, Wavevector ($\mathbf{q}$) dependence of the magnetic fluctuations along the middle (**b**) and lower (**c**) trajectory in the inset of **a** for fixed $E = 0.035$ meV. $\chi'' = S/(1 - e^{-E/k_B T})$ is shown, where S represents the intensity of the scattered neutrons. The non-magnetic background, derived from scans with q parallel to the easy moment direction c , has been subtracted. Different colours represent different temperatures. Lines correspond to equation (7) with $H = 0$ and a temperature independent $\theta(\mathbf{q})^\alpha$ expanded in even powers of $\mathbf{q}$.

impulse response is an exponential decay, $\chi(\mathbf{q}, t) \propto \exp[-(\Gamma_q/\hbar)t]$, where $\Gamma_q/\hbar$ is the magnetic decay rate, for which

$$\chi(\mathbf{q}, E) = A/(\Gamma_q - iE) \quad (1)$$

where A is a constant. The exponential form is typical for paramagnetic rare-earth insulators where the magnetization is not conserved. To understand how Γ_q can depend on temperature, we recall that the zero-energy response is the static magnetic susceptibility, which in paramagnetic insulators follows a Curie–Weiss law $\chi(\mathbf{q}, E)|_{E=0} = C/(T + \theta(\mathbf{q}))$ with Curie constant C and $\mathbf{q}$ -dependent Weiss temperature $\theta(\mathbf{q})$. At $\mathbf{q} = 0$ this is just the uniform susceptibility, which can be directly measured using a magnetometer. Comparison with equation (1) leads to the conclusion that $\Gamma_q = a(T + \theta(\mathbf{q}))$ where $a = A/C$. Hence, in paramagnetic insulators, the inverse magnetic response function

$$\chi^{-1}(\mathbf{q}, E, T) = \chi_0^{-1}(E, T) + \theta^\alpha(\mathbf{q})/C \quad (2)$$

is a sum of a non-local, purely $\mathbf{q}$ -dependent piece $\theta^\alpha(\mathbf{q})/C$ (here with $\alpha = 1$) and a local ($\mathbf{q}$ -independent) driving term $\chi_0^{-1} = (T - iE/a)/C$. Thus T and iE/a are interchangeable in expressions for $\chi(\mathbf{q}, E)$. At critical wavevectors, that is, the points in the Brillouin zone where $\theta(\mathbf{q}) = 0$, $\chi = \chi_0$ exhibits E/T scaling, meaning that

$$\chi_0 = T^{-\alpha} g(E/T) = E^{-\alpha} j(T/E) \quad (3)$$

where $j(1/y) = y^\alpha g(y)$. In the simple Curie–Weiss example, $\alpha = 1$ and $g(y) = g_0(y) = C/(1 - iy/a)$.

Although E/T -scaling at critical wavevectors is well-known and established^{7,13,17,26}, it does not capture the general dependence of χ on wavevector, magnetic field, and zero-temperature tuning parameter (for example, pressure or composition) for the quantum phase transition. However, inspection of the Curie–Weiss example suggests a set of hitherto untested scaling laws and data replotting procedures. The first law follows from insertion of the $T \rightarrow 0$ limit of equation (3) into equation (2). The result displays $E/\theta(\mathbf{q})$ scaling, that is:

$$\chi(T \rightarrow 0) = \theta(\mathbf{q})^{-\alpha} G(E/\theta(\mathbf{q})) \quad (4)$$

In the simple Curie–Weiss case $G(\delta) = g(\delta)$ (with $\delta = E/\theta$), which may be understood if we regard excursions in wavenumber or quantum control parameter as excursions in an effective Curie temperature. Letting $E \rightarrow 0$ instead of $T \rightarrow 0$ in equation (3) and

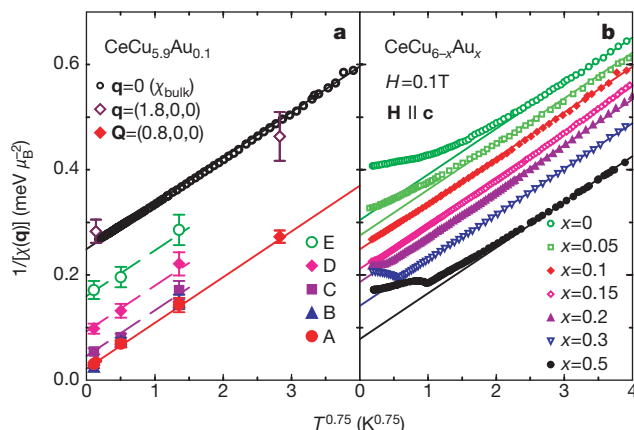


Figure 3 Temperature dependence of the inverse static susceptibility. **a**, $1/\chi(\mathbf{q})$ (from direct magnetization ($\mathbf{q} = 0$) and neutron data ($\mathbf{q} \neq 0$)), showing the same anomalous power-law dependence on temperature, over a wide range of different momentum values (labelled by letters in Fig. 2a inset). **b**, Checking the same anomalous temperature power-law of $1/\chi(\mathbf{q} = 0)$ for different Au concentrations x . The temperature below which deviations occur increases with distance to the critical concentration $x = 0.1$.

again inserting the result into equation (2) gives a second verifiable relation:

$$\chi^{-1}(\mathbf{q}, T) = (T^\alpha + \theta(\mathbf{q})^\alpha)/C \quad (5)$$

Equation (5) implies that the traditional Curie–Weiss plot, where $1/\chi$ is a straight line as a function of T , can be replaced across a series of experiments where the Weiss temperature is varied—for example, by varying $\mathbf{q}$ or composition x —by a plot where $1/\chi$ as a function of T^α is a straight line with intercept $-\theta(\mathbf{q})^\alpha$.

Finally, combining with equation (2) the knowledge that for local moment systems, the magnetic response is governed by the ratio of the Zeeman energy $g\mu_B H$ to the thermal energy $k_B T$, we obtain another scaling law:

$$(\chi(H, T)^{-1} - \theta(0)^\alpha/C)^{-1} T^\alpha = f(H/T) \quad (6)$$

This law can be tested with great sensitivity in the long-wavelength ($q = 0$) limit by using bulk magnetometry.

For systems in which the critical magnetic degrees of freedom derive from spin-density fluctuations of a Fermi sea, matters are much more complicated and the magnetic response function will not satisfy equations (2)–(6). Here we expect that non-local terms, derived from the band-structure, should depend on both E and T . At low T , the first correction to the $T = 0$ Pauli susceptibility is of order T^2 rather than T^α , as suggested by the generalized Curie–Weiss

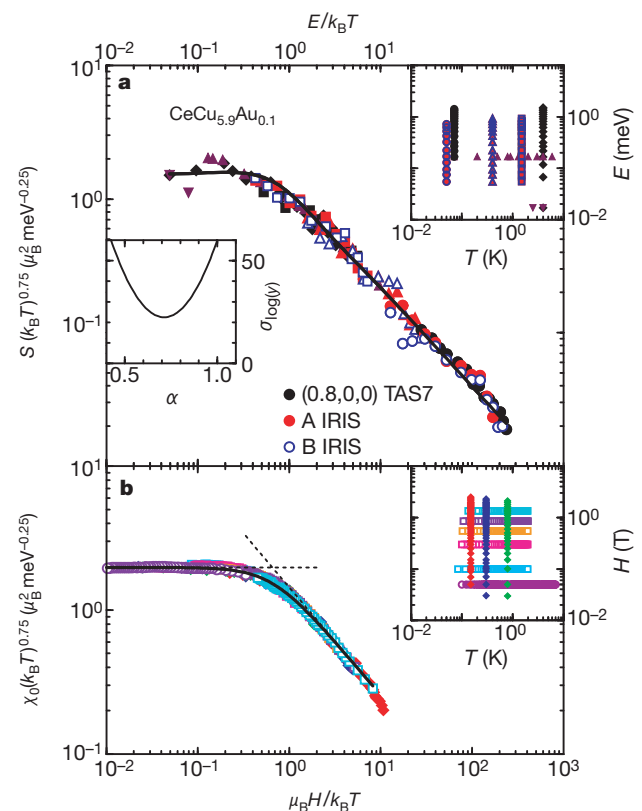


Figure 4 Temperature is the only scale for energy and field-dependent response at quantum criticality. **a**, E/T scaling plots taken at various critical $\mathbf{q}$ vectors (labelled as A and B as in Fig. 2a inset), combining previous triple-axis data with the much more extensive time-of-flight data taken at IRIS. The line is the product of the Bose–Einstein factor $[1 - \exp(-E/k_B T)]$ multiplied by the scaling function $g(y) = C/(1 - iy/a)^\alpha$ with $y = E/T$, scaling factor $a = 0.8k_B$ and exponent $\alpha = 0.75$. Right inset shows the wide range of E – T covered. Left inset shows the ‘scatter’ of the scaling plot as a function of α , being minimal for $\alpha = 0.72 \pm 0.05$. **b**, H/T -scaling of the local contribution to the uniform magnetization $M(T, H)$. $1/\chi_0 = (dM/dH)^{-1} - (4.1\mu_B^2 \text{ meV})^{-1}$. Solid line corresponds to scaling function $f(h) = (1 + h^2)^{-\alpha/2}$, with $h = g\mu_B H/k_B T$ with $\alpha = 0.75$ and an effective moment $\mu = g\mu_B = 1.5\mu_B$. Inset shows the H – T range where data were collected and scaling applies.

Ansatz. Also, T and iE are not interchangeable in the temperature-dependent susceptibility. In particular, above two dimensions⁴, the quantum critical fluctuations of a spin density wave are non-interacting at low energies and long wavelengths, in which case E/T scaling⁷ never occurs. Likewise, none of the less-familiar forms—equations (4)–(6)—accounting for the $\mathbf{q}$ -, T -, x and H -dependence should obtain. Of special significance is that an external magnetic field modifies the band structure, implying an influence on the non-local terms and an inability to scale the field-dependent effects simply in terms of the ratio $g\mu_B H/k_B T$ (ref. 29).

We have thus reduced the problem of discriminating between the pictures of quantum criticality shown in Fig. 1b and c to discovering two things. First, whether the largely untested laws, given in equations (4)–(6) describe the H -, T -, E - and x -dependent magnetic response of $\text{CeCu}_{6-x}\text{Au}_x$, and second, verifying that with an order of magnitude improvement in energy range, E/T -scaling obtains at the critical composition and wavevector. To perform this programme, we used a low-temperature d.c. magnetometer²⁷ and the IRIS spectrometer²⁸ at the ISIS spallation source. IRIS represents a substantial improvement over the instruments previously used to examine the magnetic fluctuations in $\text{CeCu}_{5.9}\text{Au}_{0.1}$: first, it surveys the magnetic fluctuation spectrum over a very wide range of wavenumbers; second, it has an energy resolution of $9\text{ }\mu\text{eV} \equiv 0.1\text{ K}$ —almost ten times better than that of previous experiments; and third, it yields data in absolute units established using the simultaneously measured elastic incoherent scattering from the sample.

As x is varied in $\text{CeCu}_{6-x}\text{Au}_x$, Bragg reflections due to antiferromagnetism appear at specific wavenumbers on a ‘butterfly-shaped’ critical line¹⁶ (black crosses in Fig. 2a). At the quantum critical concentration $x = x_c$, the Bragg peaks are replaced by magnetic critical scattering—peaked on the same ‘butterfly’¹⁵. Figure 2b and c shows scans through the ‘butterfly’, obtained for the very small energy $E = 35\text{ }\mu\text{eV}$. On warming from 50 mK to 0.4 K ($\approx E/k_B$), the magnetic response at the butterfly is somewhat reduced, while on going to 1.5 K $= 4E/k_B$, it is greatly suppressed, showing that when T crosses E/k_B , the magnetic response begins to change rapidly. Figure 2a shows energy scans collected at 50 mK at the various (nearly) fixed momentum transfers indicated in the inset, and plotted following the scaling given by equation (4); we used the exponent $\alpha \approx 0.75$ obtained from the E/T scaling found in our previous critical wavenumber ($\mathbf{q} = (1.2, 0, 0)$) experiment¹⁷, which is verified with much greater confidence below. That the data collapse onto a single curve is clearly consistent with equation (4), but there is sufficient scatter to necessitate a check of the other untested law (equation (5)) before we can claim to have verified the local moment model giving rise to equation (2).

Figure 3 shows the modified Curie–Weiss plots suggested by equation (5) for $\text{CeCu}_{6-x}\text{Au}_x$, where the effective Weiss temperature is varied either by changing q in the neutron scattering experiment, or by changing x in the bulk (low field) magnetometry. Near the QCP, the static susceptibility follows the modified Curie–Weiss law over the entire two-decade temperature range of our measurements. As x moves away from x_c in either direction, $\chi^{-1}(T)$ becomes more curved as $T \rightarrow 0$, although there still is a quantum critical regime bounded by a cross-over temperature T^* , which increases with distance from x_c ; T^* is defined so that for $T > T^*$ the T^α power law seems to hold (Fig. 3b), while for $T < T^*$ a tendency towards less singular (heavy) Fermi liquid behaviour is seen. Figure 3a also shows that at quantum criticality ($x \approx 0.1$) the same law is followed for general $\mathbf{q}$, where $\chi'(\mathbf{q}) = \chi'(\mathbf{q}, E = 0)$ is obtained via the Kramers–Kronig relation from the neutron data. Indeed, the parallelism of the lines in Fig. 3a shows not only that a single exponent α is sufficient to account for all of the magnetometry and neutron data, but also that the Curie constant C is—within the limited statistical accuracy of the latter—the same for all wavenumbers. The overlap between the neutron data at $\mathbf{q} = 0$ (the bulk

results) and $\mathbf{q} = (1.8, 0, 0)$ (far from the ‘butterfly’ but close to $(2, 0, 0)$, equivalent to $(0, 0, 0)$) is reassuring, and shows that $\mathbf{q} = 0$ is a typical non-critical wavenumber, with a spectrum likely to resemble that measured at $\mathbf{q} = (1.8, 0, 0)$. A final result that can be extracted from Fig. 3 is another estimate of the wavenumber-dependent Weiss temperature $\theta(\mathbf{q})$, which is simply the x -axis intercept, taken to the power $1/\alpha$, of the lines going through the data for each $\mathbf{q}$. The outcome is entirely consistent with that of the $E/\theta(\mathbf{q})$ scaling analysis in Fig. 2a.

We now turn to the energy- and temperature-dependent spectroscopies with the best signal-to-noise ratios: the neutron measurements of the scattering with $\mathbf{q}$ on the butterfly, and the H -dependent static magnetization. In Fig. 4 we show the former, plotted versus E/T , collected at a wavenumber $(0.8, 0, 0)$. This data set, which densely covers four decades of E/T , confirms that the scaling is optimal (see left inset) for $\alpha = 0.75 \pm 0.05$. Figure 4b shows the collapse of the bulk magnetization data as suggested by the H/T scaling form, equation (6). The similarities in the scaling function and exponent ($\alpha = 3/4$), obtained in this entirely independent assay, to those obtained via neutron measurements show that the magnetic response is driven by a singular local term. Two details emerge. First, the effective moment (defined in Fig. 4 legend) $\mu = g\mu_B = 1.5\mu_B$ is of the same order as the atomic moment ($0.6\mu_B$) estimated for the q -space average of the neutron data integrated to $E = 1\text{ meV}$. This rules out large, randomly occurring ferromagnetic clusters as a vital feature of the QCP³⁰. Second, the scaling function $f(h)$ (where $h = g\mu_B H/k_B T$) has the power-law asymptote $h^{-\alpha}$ at large h , which decays much more slowly than the $\exp(-h)$ asymptote derived from the Brillouin function. Thus, the property responsible for the critical behaviour is an atomically local magnetic moment, with a size comparable with a single spin $S = 1/2$, but also with an intrinsically critical response to an external field.

Our experiments demonstrate that the magnetic fluctuations in the heavy-fermion metal $\text{CeCu}_{6-x}\text{Au}_x$ are very simple near the QCP at $x = 0.1$. Not only do our high-resolution neutron data satisfy E/T scaling at the critical wavenumbers with much greater certainty than before, but, together with magnetization data, they satisfy two hitherto untested scaling laws and a generalized Curie–Weiss model, all of which emerge from a local-moment model for the QCP. Detailed quantitative analysis (see figure legends) reveals that an algebraic form

$$\chi^{-1}(\mathbf{q}, E, T) = \left[\left(T^2 + \left(\frac{g\mu_B}{k_B} \right)^2 H^2 \right)^{1/2} - iE/a \right]^\alpha + \theta(\mathbf{q})^\alpha \Big] C^{-1} \quad (7)$$

$$= \chi_0^{-1}(E, T) + \theta(\mathbf{q})^\alpha / C$$

accounts for all of our H -, $\mathbf{q}$ - and T -dependent results. Equation (7) is the most straightforward generalization of the Curie–Weiss description of a critical point, and has only four parameters—an overall amplitude, a scaling factor $a \approx k_B$, an exponent α and the function $\theta(\mathbf{q})$.

So in answer to the question posed earlier—whether an itinerant (Fig. 1b) or local (Fig. 1c) picture of the magnetic QCP applied to this heavy-fermion system—it is the latter that is appropriate. We also note the appearance of a single unusual exponent $\alpha = 0.75$ in the scaling properties at widely different wavevectors. The E/T and H/T scaling with the unusual exponent $\alpha = 0.75 < 1$ mean that the magnetic properties of $\text{CeCu}_{6-x}\text{Au}_x$ are more severally non-analytic than those of both Fermi liquids and conventional insulating magnets.

We conclude that a Fermi liquid, CeCu_6 , with one of the largest known effective carrier masses, is remarkably close to an instability marked by a new type of critical behaviour. The moments associated with the tightly bound f electrons are on the brink of separation from the conduction electrons—how close is apparent from the

approach of the inverse magnetic susceptibility to the anomalous T^α asymptote as T rises above 2 K (Fig. 3b). The locality of the fluctuations at the quantum critical point suggests the development of an unusual kind of localized excitation—derived from the ambiguity of whether the f electrons are counted within the Fermi sea or not. These excitations are neither heavy quasi-particles nor conventional spin excitations, and might best be regarded as Fermi surface ‘shredders’: degrees of freedom that are ultimately responsible for both the non-Fermi-liquid behaviour and the unusual local scaling behaviour at the quantum critical point. □

Received 18 April; accepted 26 July 2000.

- Anderson, P. W. New approach to the theory of superexchange interactions. *Phys. Rev.* **115**, 2–12 (1959).
- Stoner, E. C. *Proc. R. Soc. Lond. A* **154**, 656–678 (1936).
- Overhauser, A. W. New mechanism of antiferromagnetism. *Phys. Rev. Lett.* **3**, 414–415 (1959).
- Hertz, J. A. Quantum critical phenomena. *Phys. Rev. B* **14**, 1165–1184 (1976).
- Millis, A. J. Effect of a nonzero temperature on quantum critical points in itinerant ferion systems. *Phys. Rev. B* **48**, 7183–7196 (1993).
- Chakravarty, S., Halperin, B. I. & Nelson, D. R. Low temperature behavior of a 2D quantum antiferromagnet. *Phys. Rev. Lett.* **60**, 1057–1060 (1988).
- Sachdev, S. & Ye, J. Universal quantum-critical dynamics of two-dimensional antiferromagnets. *Phys. Rev. Lett.* **69**, 2411–2414 (1992).
- Stewart, G. R., Fisk, Z. & Wire, M. S. New Ce heavy-fermion system: CeCu_6 . *Phys. Rev. B* **30**, 482–484 (1984).
- Onuki, Y., Shimizu, Y. & Komatsubara, T. Magnetic property of a new Kondo lattice intermetallic compound: CeCu_6 . *J. Phys. Soc. Jpn* **53**, 1210–1213 (1984).
- Doniach, S. The Kondo lattice and weak antiferromagnetism. *Physica B* **91**, 231–234 (1977).
- Aeppli, G. & Broholm, C. in *Handbook on the Physics and Chemistry of Rare Earths* Vol. 19 (ed. Gschneidner, K. A. Jr et al.) 123–175 (Elsevier, Amsterdam, 1994).
- Mathur, N. D. et al. Magnetically mediated superconductivity in heavy fermion compounds. *Nature* **394**, 39–43 (1998).
- Varma, C. M., Littlewood, P. B., Schmitt-Rink, S., Abrahams, E. & Ruckenstein, A. E. Phenomenology of the normal state of Cu-O high temperature superconductors. *Phys. Rev. Lett.* **63**, 1996–1999 (1989).
- Löhneysen, H.v. et al. Non-Fermi-liquid behavior in a heavy-fermion alloy at a magnetic instability. *Phys. Rev. Lett.* **72**, 3262–3265 (1994).
- Stockert, O., Löhneysen, H.v., Rosch, A., Pyka, N. & Loewenhaupt, M. Two-dimensional fluctuations at the quantum-critical point of $\text{CeCu}_{6-x}\text{Au}_x$. *Phys. Rev. Lett.* **80**, 5627–5630 (1998).
- Löhneysen, H.v. et al. Magnetic order and transport in the heavy-fermion system $\text{CeCu}_{6-x}\text{Au}_x$. *Eur. Phys. J. B* **5**, 447–456 (1998).
- Schröder, A., Aeppli, G., Bucher, E., Ramazashvili, R. & Coleman, P. Scaling of magnetic fluctuations near a quantum phase transition. *Phys. Rev. Lett.* **80**, 5623–5626 (1998).
- Rosch, A., Schröder, A., Stockert, O. & Löhneysen, H.v. Mechanism for the non-Fermi-liquid behavior in $\text{CeCu}_{6-x}\text{Au}_x$. *Phys. Rev. Lett.* **79**, 159–162 (1997).
- Ruck, M. et al. Structure and electrical resistivity of the heavy fermion compound CeCu_2Au . *Acta Crystallogr. B* **49**, 936–941 (1993).
- Schröder, A., Lynn, J. W., Erwin, R. W., Loewenhaupt, M. & Löhneysen, H.v. Magnetic structure of the heavy fermion alloy $\text{CeCu}_{5.5}\text{Au}_{0.5}$. *Physica B* **199–200**, 47–48 (1994).
- Stockert, O. et al. Incommensurate antiferromagnetism and magnetic correlations in $\text{CeCu}_{6-x}\text{Au}_x$. *Physica B* **230–232**, 247–249 (1997).
- Si, Q., Smith, J. L. & Ingersent, K. Quantum critical behavior in Kondo systems. *Int. J. Mod. Phys. B* **13**, 2331–2342 (1999).
- Rosch, A. Interplay of disorder and spin fluctuations in the resistivity near a quantum critical point. *Phys. Rev. Lett.* **82**, 4280–4283 (1999).
- Doniach, S. in *Proc 5th Int. Conf. on Valence Fluctuations* 179–185 (Plenum, New York, 1987).
- Moriya, T. & Takimoto, T. Anomalous properties around magnetic instability in heavy electron systems. *J. Phys. Soc. Jpn* **64**, 960–969 (1995).
- Aronson, M. C. et al. Non-Fermi-liquid scaling of the magnetic response in $\text{UCu}_{5-x}\text{Pd}_x$ ($x = 1.1, 1.5$). *Phys. Rev. Lett.* **75**, 725–728 (1995).
- Schröder, A., Schlager, H. G. & Löhneysen, H.v. Magnetization of CeCu_6 at low temperatures. *J. Magn. Mater.* **108**, 47–49 (1992).
- Carille, C. J. & Adams, M. A. The design of the IRIS inelastic neutron spectrometer and improvements to its analysers. *Physica B* **182**, 431–440 (1992).
- Sachdev, S. Quantum phase transitions and conserved charges. *Z. Phys. B* **94**, 469–479 (1994).
- Castro Neto, A. H., Castilla, G. & Jones, B. A. Non-Fermi liquid behavior and Griffiths phase in f -electron compounds. *Phys. Rev. Lett.* **81**, 3531–3534 (1996).

Acknowledgements

A.S., O.S. and H.v.L. were supported by the Deutsche Forschungsgemeinschaft; R.C. was supported by the Oak Ridge National Laboratory Postdoctoral Research Associates program, administered jointly by the Oak Ridge National Laboratory and the Oak Ridge Institute for Science and Education; P.C. and R.R. were supported by the National Science Foundation. R.R. acknowledges partial support from the MacArthur Chair endowed by the John D. & Catherine T. MacArthur Foundation at the University of Illinois.

Correspondence and requests for materials should be addressed to P.C. (e-mail: Coleman@physics.rutgers.edu).

Detection of geometric phases in superconducting nanocircuits

Giuseppe Falci*, Rosario Fazio*, G. Massimo Palma†§, Jens Siewert*§ & Vlatko Vedral‡

* Dipartimento di Metodologie Fisiche e Chimiche (DMFCI), Università di Catania, viale A. Doria 6, I-95125 Catania, Italy

† Dipartimento di Scienze Fisiche ed Astronomiche (DSFA), Università di Palermo, via Archirafi 36, I-90123 Palermo, Italy

‡ Centre for Quantum Computation, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, UK

§ Istituto Nazionale per la Fisica della Materia (INFN), Unità di Catania e Palermo, Italy

When a quantum-mechanical system undergoes an adiabatic cyclic evolution, it acquires a geometrical phase factor¹ in addition to the dynamical one; this effect has been demonstrated in a variety of microscopic systems². Advances in nanotechnology should enable the laws of quantum dynamics to be tested at the macroscopic level³, by providing controllable artificial two-level systems (for example, in quantum dots⁴ and superconducting devices^{5,6}). Here we propose an experimental method to detect geometric phases in a superconducting device. The setup is a Josephson junction nanocircuit consisting of a superconducting electron box. We discuss how interferometry based on geometrical phases may be realized, and show how the effect may be applied to the design of gates for quantum computation.

Interferometry based on Berry's phase has been proposed to realize quantum gates. This is at the heart of geometric quantum computation that has been implemented in NMR⁷. Quantum computation based on non-abelian phases has been proposed⁸. Quantum computers can be viewed as programmable quantum interferometers⁹. Initially prepared in a superposition of all the possible input states, the computation evolves in parallel along all its possible paths, which will interfere constructively towards the desired output state. This intrinsic parallelism in the evolution of a quantum system allows for exponentially more efficient ways of performing computation¹⁰.

The quest for large-scale integrability has stimulated an increasing interest in superconducting nanocircuits^{11–16} as possible candidates for the implementation of a quantum computer. Mesoscopic Josephson junctions can be prepared in a controlled superposition of charge states^{17,18} and the coherent time evolution in a Josephson charge qubit has been recently observed⁵. These are the first important experimental steps towards the implementation of a solid-state quantum computer.

An important aspect of this research activity is that it unveils fundamental problems related to the quantum mechanical behaviour of macroscopic systems. Here we study quantum interferometry based on geometric phases in mesoscopic Josephson devices. A well known example in classical superconductivity is the Aharonov–Bohm (AB) effect which manifests itself, for example, in the periodic modulation of the critical current in classical SQUIDS (ref. 19). The AB effect provides evidence of the macroscopic coherence of the superconducting condensate. We will show that it is possible to detect geometric phases other than the AB phase in the coherent dynamics of superconducting nanocircuits. Quantum interferometry based on these geometric phases allows us to develop a new design of gates for quantum computation using charge qubits (the designs proposed up to now for superconducting quantum gates use interferometry of dynamical phases). Indeed, it is possible to introduce conditional geometric phases, that is, in which the geometric phase on one qubit is controlled by the state of the

neighbouring one. Such a feature, in itself a remarkable effect in the theory of geometric phases, is an important requirement of quantum computation.

The setup we consider is shown in Fig. 1a. It consists of a superconducting electron box formed by an asymmetric SQUID, pierced by a magnetic flux Φ and with an applied gate voltage V_x . The device operates in the charging regime, that is, the Josephson couplings $E_{J1(2)}$ of the junctions are much smaller than the charging energy E_{ch} . We further assume that the temperature is much lower than $E_{J1(2)}$. The hamiltonian is¹⁹

$$H = E_{ch}(n - n_x)^2 - E_J(\Phi) \cos(\theta - \alpha) \quad (1)$$

$$E_J(\Phi) = \sqrt{(E_{J1} - E_{J2})^2 + 4E_{J1}E_{J2} \cos^2\left(\pi \frac{\Phi}{\Phi_0}\right)}$$

$$\tan \alpha = \frac{E_{J1} - E_{J2}}{E_{J1} + E_{J2}} \tan\left(\pi \frac{\Phi}{\Phi_0}\right)$$

where and $\Phi_0 = h/2e$ is the (superconducting) quantum of flux. The phase difference across the junction θ and the number of Cooper pairs n are canonically conjugate variables $[\theta, n] = i$. The parameters of the hamiltonian can be controlled. The offset charge $2en_x$ can be tuned by changing V_x (see Fig. 1a) and the coupling $E_J(\Phi)$ depends on Φ , as in the device proposed in refs 11 and 12. We propose the use of an asymmetric SQUID, which permits control of the phase shift $\alpha(\Phi)$ as well.

At temperatures much lower than E_{ch} , if n_x varies around the value 1/2, only the two charge eigenstates $n = 0, 1$ are important. They constitute the basis $\{|0\rangle, |1\rangle\}$ of the computational Hilbert space of the qubit^{11–13}. The effective hamiltonian is obtained by projecting equation (1) on the computational Hilbert space, and reads $H_B = -(1/2)\mathbf{B} \cdot \boldsymbol{\sigma}$, where we have defined the fictitious field $\mathbf{B} \equiv (E_J \cos \alpha, -E_J \sin \alpha, E_{ch}(1 - 2n_x))$ and $\boldsymbol{\sigma}$ are the Pauli matrices.

The system thus behaves like an artificial spin in a magnetic field. Charging couples the system to B_z whereas the Josephson term determines the projection in the xy plane. By changing V_x and Φ the qubit hamiltonian H_B describes a cylindroid in the parameter space $\{\mathbf{B}\}$. The presence of higher charge states leads to quantum leakage.

This effect can be minimized by a fine-tuning of the parameters of the devices¹⁶.

We now consider the adiabatic time evolution of the qubit hamiltonian. By changing H_B around a circuit in the parameter space $\{\mathbf{B}\}$, the eigenstates will accumulate a Berry phase proportional to the solid angle that the circuit subtends at the degeneracy point, $\mathbf{B} = 0$. The Berry phase $\gamma_B(\Phi_M, n_{xm})$ accumulated by the two lowest eigenstates along a circuit where n_x is varied from n_{xm} to 1/2 and the flux from zero to Φ_M (see the inset of Fig. 1a) is plotted in Fig. 2.

We note that non-trivial loops with a controllable Berry phase are possible in our setup thanks to the asymmetry in the SQUID. In the symmetric case $\alpha(\Phi) = 0$ and the evolution takes place in a strip of the plane $B_y = 0$; that is, the Berry phase is zero. The distinct feature of the (geometric) interferometry proposed here is that the geometric phase does depend on both gate voltage and flux. It differs from 'classical' interference in mesoscopic junctions, achieved either by changing the magnetic field (in the flux regime) or by a gate voltage modulation (in the charge regime).

The experimental setup required to measure $\gamma_B(\Phi_M, n_{xm})$ is already available and corresponds to that of ref. 5. We now describe a procedure to measure the Berry phase. The system is prepared in the ground state of the hamiltonian at $n_x = 0$ and $\Phi = 0$ and then a sudden switching to the point $n_x = 1/2$ and $\Phi = 0$ ($\mathbf{B}_0 = (E_J(0), 0, 0)$) is applied to produce the initial hamiltonian H_{B0} . The initial charge state is then a linear superposition of the eigenstates $\{|\mathbf{B}_{0+}\rangle, |\mathbf{B}_{0-}\rangle\}$ of the hamiltonian H_{B0} . This is analogous to the splitting of the photon wavefunction at the first beamsplitter of a Mach-Zender interferometer. A phase difference between the states $|\mathbf{B}_{0+/-}\rangle$ can be introduced by adiabatically dragging H_B along a closed loop (see the inset of Fig. 1a). The phases acquired in this way will have both a geometrical and a dynamical component. In order to eliminate the latter it is sufficient to perform a NOT gate, for instance by applying an a.c. gate voltage pulse that swaps the eigenstates. If the same loop as before is covered backwards by H_B , the geometrical component of the phase will add and the dynamical phases will cancel each other⁷. The final step is to measure the charge state of the qubit. The probability of measuring a charge $2e$ ($n = 1$) in the box at the end of this procedure—up to negligible terms of the order of

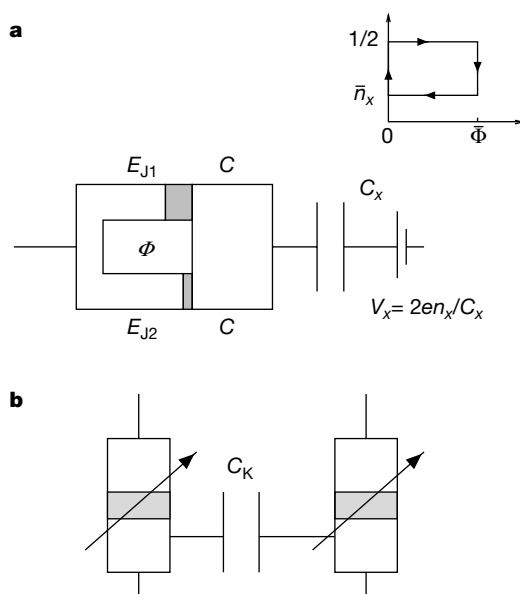


Figure 1 Schematic design of the system and the gates. **a**, The system consists of a superconducting electron box formed by an asymmetric SQUID, pierced by a magnetic flux Φ , and with an applied gate voltage V_x . The offset charge n_x defined in the paper is proportional to V_x . The device operates in the charge regime, that is, $E_{J1}, E_{J2} \ll E_{ch}$. The

cyclic evolution of the hamiltonian is obtained by changing the gate voltage and the magnetic flux along the path shown in the inset. **b**, Schematic design of the two-bit gate. The two qubits are coupled capacitively.

$(E_j(\Phi_0/2)/E_{ch})^2$ —is given by:

$$P(1) = \sin^2(2\gamma_B) \quad (2)$$

independently on the time elapsed. The result given in equation (2) holds as long as the second adiabatic loop retraces exactly the first one, in order to cancel the dynamical phase. If this is not the case, the measurement will be affected by random errors that will reduce the amplitude of the interference fringes. The dephasing time should be longer than the time needed for the adiabatic manipulation, which has to be longer than the typical timescale for the dynamics, $\hbar/|E_{j1} + E_{j2}|$. The results of ref. 5 indicate that with the present technology there is already sufficient control of the dynamics to detect the geometric phase.

Geometric interferometry in mesoscopic superconducting devices can be conveniently discussed in the language of quantum computation. Conditional geometric interferometry, which corresponds to the implementation of a universal two-qubit gate is realized by coupling capacitively two asymmetric SQUIDS (see Fig. 1b). This provides a framework for the implementation of geometric quantum computation in a solid-state device. If the coupling capacitance C_K is smaller than the others, the hamiltonian is

$$H_G = \sum_{i=1,2} H_i + \delta E(n_1 - n_{x,1})(n_2 - n_{x,2}) \quad (3)$$

where H_i refer to the two uncoupled qubits defined in equation (1) and $\delta E = 2E_{ch}C_K/C$ is the charging energy which derives from the capacitive coupling. Gate voltages and magnetic fluxes can be independently fixed for the two qubits. This setup can be used to implement a controlled phase shift; that is, the geometric phase of one qubit (the target) depends on the state of the neighbouring one

(the control qubit) and to detect it. Because of universality in quantum computation²¹, this gate can be used to perform any computational task, if complemented with the single qubit operations described above. The gate voltage of the control qubit is kept away from the resonance condition. As a result of the coupling, the effective charging of the target qubit, and consequently the surface spanned in the parameter space, depends on the state of the control qubit. This implies that the Berry phase γ_B is conditional. In the basis of the charge states, the unitary operator that describes this gate is

$$\begin{pmatrix} \exp(-i\gamma_0) & 0 & 0 & 0 \\ 0 & \exp(i\gamma_0) & 0 & 0 \\ 0 & 0 & \exp(-i\gamma_1) & 0 \\ 0 & 0 & 0 & \exp(i\gamma_1) \end{pmatrix}$$

where γ_i is the Berry phase when the charge state of the control qubit is $i = 0, 1$. Measurement of the conditional phase shift can be performed by following the procedure outlined previously. One important issue is that the NOT operation needed to cancel the dynamical phase can be performed with an a.c. voltage bias pulse whose amplitude does not depend on the value of the control qubit. As it is not possible to switch off the Josephson coupling of the control qubit completely, the real gate will have very small off-diagonal terms that are irrelevant on the timescale of manipulation.

Quantum interferometry provides a clear conceptual framework to describe in a unified language a broad range of quantum phenomena and to transfer ideas for new experimental investigation of fundamental aspects of quantum theory. We have described a new class of quantum geometric phase effects in small Josephson junctions, other than conventional AB effect. Adiabatic pumping of Cooper pairs was interpreted in terms of the Berry phase in ref. 22. We discuss how to realize interferometry based on geometrical phases, and show how this effect can find application in the design of new gates for quantum computation. The unifying concept underlying these two apparently distinct topics is the fact that both the detection procedure of geometric phases and quantum computation can be described in interferometric terms. $\square$

Received 23 May; accepted 25 July 2000.

- Berry, M. V. Quantal phase factors accompanying adiabatic changes. *Proc. R. Soc. Lond. A* **392**, 45–57 (1984).
- Shapere, A. & Wilczek, F. (eds) *Geometric Phases In Physics* (World Scientific, Singapore, 1989).
- Caldeira, A. O. & Leggett, A. J. Influence of dissipation on quantum tunneling in macroscopic systems. *Phys. Rev. Lett.* **46**, 211–214 (1981).
- Oosterkamp, T. H. *et al.* Microwave spectroscopy of a quantum dot molecule. *Nature* **395**, 873–876 (1998).
- Nakamura, Y., Pashkin, Yu. & Tsai, J. S. Coherent control of macroscopic quantum states in a single-Cooper-pair box. *Nature* **398**, 786–788 (1999).
- Friedman, J. R., Patel, V., Chen, W., Tolpygo, S. K. & Lukens, J. K. Quantum superposition of distinct macroscopic states. *Nature* **406**, 43–46 (2000).
- Jones, J., Vedral, V., Ekert, A. K. & Castagnoli, C. Geometric quantum computation using nuclear magnetic resonance. *Nature* **403**, 869–871 (2000).
- Zanardi, P. & Rasetti, M. Holonomic quantum computation. *Phys. Lett. A* **264**, 94–99 (1999).
- Cleve, R., Ekert, A. K., Henderson, L., Macchiavello, C. & Mosca, M. On quantum algorithms. *Complexity* **4**, 33–42 (1998).
- Ekert, A. & Jozsa, R. Quantum computation and Shor's factoring algorithm. *Rev. Mod. Phys.* **68**, 733–753 (1996).
- Shnirman, A., Schön, G. & Hermon, Z. Quantum manipulation of small Josephson junctions. *Phys. Rev. Lett.* **79**, 2371–2374 (1997).
- Makhlin, Y., Schön, G. & Shnirman, A. Josephson-junction qubits with controlled couplings. *Nature* **398**, 305–307 (1999).
- Averin, D. V. Josephson-junction qubits with controlled couplings. *Solid State Commun.* **105**, 659–664 (1998).
- Mooij, J. E. *et al.* Josephson persistent current qubit. *Science* **285**, 1036–1039 (1999).
- Ioffe, L. B., Geshkenbein, V. B., Feigelman, M. V., Fauchere, A. L. & Blatter, G. Environmentally decoupled *sds*-wave Josephson junctions for quantum computing. *Nature* **398**, 679–681 (1999).
- Fazio, R., Palma, G. M. & Siewert, J. Fidelity and leakage of Josephson qubits. *Phys. Rev. Lett.* **83**, 5385–5388 (1999).
- Matters, M., Elion, W. & Mooij, J. E. Influence of controlled quantum-mechanical charge and phase fluctuations on Josephson tunneling. *Phys. Rev. Lett.* **75**, 721–724 (1995).
- Bouchiat, V., Vion, D., Joyez, P., Esteve, D. & Devoret, M. Quantum coherence with a single Cooper pair. *Phys. Scr.* **T76**, 165–170 (1998).
- Tinkham, M. *Introduction to Superconductivity* 2nd edn (McGraw-Hill, New York, 1996).

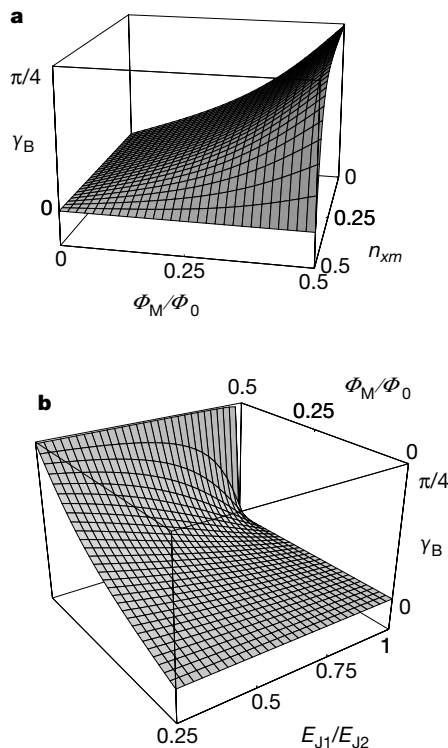


Figure 2 Geometric phase produced by the adiabatic manipulations. **a**, The Berry phase is calculated for the loop shown in the inset of Fig. 1a as a function of the flux and the external charge: $\gamma_B(\Phi_M, n_{xm}) = \pm [E_{ch}(1 - 2n_{xm})/4(E_{j1}E_{j2})^{1/2}] \lambda \mu \Pi(\pi\Phi_M/\Phi_0, 1 - \mu^2, \lambda)$ where $\Pi(x, n, k)$ is the elliptic function of the third kind²⁰, $\lambda^2 = 4E_{j1}E_{j2}/\{[E_{ch}(1 - 2n_{xm})]^2 + (E_{j1} + E_{j2})^2\}$ and $\mu = (E_{j1} - E_{j2})/(E_{j1} + E_{j2})$. In this plot the parameters are $E_{j1} = 0.25E_{j2}$ and $E_{ch} = 5(E_{j1} + E_{j2})$. **b**, The Berry phase is plotted as a function of the asymmetry parameter E_{j1}/E_{j2} and the flux for the path where $n_{xm} = 0$.

20. Gradshteyn, I. S. & Ryzhik, I. M. *Tables of Integrals, Series, and Products* 5th edn (Academic, London, 1999).
21. Deutsch, D., Barenco, A. & Ekert, A. Universality in quantum computation. *Proc. R. Soc. Lond. A* **449**, 669–677 (1995).
22. Pekola, J. P., Toppari, J. J., Aunola, M., Savolainen, M. T. & Averin, D. V. Adiabatic transfer of Cooper pairs in arrays of Josephson junctions. *Phys. Rev. B* **60**, R9931–R9934 (1999).

Acknowledgements

We thank D.V. Averin, A. Ekert, G. Giaquinta, J. Jones, B. Pannetier and E. Paladino for helpful discussions. G.F. acknowledges kind hospitality at LEPES-CNRS (Grenoble). Part of the work of R.F. was done at ISI-Torino. This work was supported by the European Community (TMR, IST-SQUIBIT, IST-EQUIP), by INFN-PRA-SSQI and by ELSAG S.p.A.

Correspondence and requests for materials should be addressed to G.F. (e-mail: gfalci@dmfci.ing.unict.it).

Ordering and self-organization in nanocrystalline silicon

G. F. Grom^{*}, D. J. Lockwood[†], J. P. McCaffrey[‡], H. J. Labbé[†], P. M. Fauchet^{*}, B. White Jr[‡], J. Diener[§], D. Kovalev[§], F. Koch[§] & L. Tsybeskov^{*}

^{*} Nanoscale Silicon Research Initiative, Department of Electrical and Computer Engineering, University of Rochester, Rochester, New York 14618, USA

[†] Institute for Microstructural Sciences, National Research Council, Ottawa K1A 0R6, Canada

[‡] Motorola DigitalDNA Laboratories, Austin, Texas, USA

[§] Technische Universität München, Physik-Department E16, D 85747 Garching, Germany

The spontaneous formation of organized nanocrystals in semiconductors has been observed^{1–5} during heteroepitaxial growth and chemical synthesis. The ability to fabricate size-controlled silicon nanocrystals encapsulated by insulating SiO₂ would be of significant interest to the microelectronics industry. But reproducible manufacture of such crystals is hampered by the amorphous nature of SiO₂ and the differing thermal expansion coefficients of the two materials. Previous attempts^{6–10} to fabricate Si nanocrystals failed to achieve control over their shape and crystallographic orientation, the latter property being important in systems such as Si quantum dots. Here we report the self-organization of Si nanocrystals larger than 80 Å into brick-shaped crystallites oriented along the $\langle 111 \rangle$ crystallographic direction. The nanocrystals are formed by the solid-phase crystallization of nanometre-thick layers of amorphous Si confined between SiO₂ layers. The shape and orientation of the crystallites results in relatively narrow photoluminescence, whereas isotropic particles produce qualitatively different, broad light emission. Our results should aid the development of maskless, reproducible Si nanofabrication techniques.

Sample structures were prepared in the form of nc-Si/SiO₂ superlattices¹¹. Figure 1 compares transmission electron microscope (TEM) micrographs of nc-Si superlattices with 42-Å (Fig. 1a), 85-Å (Fig. 1b) and 200-Å (Fig. 1c) thick nanocrystalline (nc)-Si layers crystallized under the same conditions. We found that a 42-Å thick nc-Si layer consists mostly of spherical and elliptical nanocrystals with significant variations in their shape (Fig. 1a). An 85-Å thick nc-Si layer consists mainly of square-shaped Si nanocrystals. For both samples the SiO₂ separating layers look almost equally planar. A superlattice with a 200-Å thick nc-Si layer contains well-defined, brick-shaped Si nanocrystals with a lateral dimension 2.2–2.5 times larger than the vertical dimension (Fig. 1c). Figure 1d shows a

detailed image of a perfectly shaped Si nanocrystal brick with a lateral size of about 500 Å and a vertical size of 200 Å. The propensity to form bricks of nc-Si is just one remarkable feature of this unusual system.

The TEM analysis in samples containing Si nanocrystals is complicated by the absence of single crystal planes over the entire sample area. Therefore, such an analysis needs to be complemented by another characterization technique, and inelastic light scattering spectroscopy is one of the most sensitive techniques to study heterointerfaces^{12–16}. The spectrum from an 85-Å nc-Si/35-Å amorphous (a)-SiO₂ superlattice with 20 periods is shown in Fig. 2a. A high spectral resolution and very low diffusive scattering permit the observation and identification of scattering from the zero-order (or Brillouin scattering) peak at $\omega_{Br} = 3.3 \text{ cm}^{-1}$ and the first doublet of superlattice folded longitudinal acoustic (FLA) phonons $m = \pm 1$ at $\omega_{-1} = 17.5 \text{ cm}^{-1}$ and $\omega_{+1} = 24.1 \text{ cm}^{-1}$. We note that $\omega_{+1} - \omega_{-1} = 2\omega_{Br}$, exactly as expected from the theory¹⁴. The $m = \pm 1$ doublet was studied using different excitation wavelengths (457.9 nm, 476.5 nm, 488.0 nm, 496.5 nm and 514.5 nm lines of the Ar⁺ laser). The Brillouin minizone^{14,15} for this sample was reconstructed using the measured $\omega_{\pm m}$ and the calculated normalized phonon wavevectors $q_p(\lambda)/q_{mz} = 2(\omega_{+1} - \omega_{-1})/(\omega_{+1} + \omega_{-1})$ (Fig. 2a, inset). The use of different wavelengths in the excitation allowed the observation of inelastic light scattering from the longitudinal acoustic phonons with normalized wavevectors in the 0.26–0.31 range, where the phonon dispersion is expected to be linear. By fitting the experimental data, we calculated the overall nc-Si/SiO₂ superlattice sound velocity as $v_{SL} = (7.5 \pm 0.2) \times 10^5 \text{ cm s}^{-1}$, and the sound velocity in 85-Å Si nanocrystals as $v_{nc-Si} = (8.7 \pm 0.5) \times 10^5 \text{ cm s}^{-1}$. The latter number is close to the sound velocity in crystalline (c)-Si along the $\langle 100 \rangle$ direction ($v_{Si(100)} = 8.5 \times 10^5 \text{ cm s}^{-1}$), and this result suggests that Si nanocrystals within nc-Si superlattices are not oriented randomly.

Raman scattering associated with FLA phonon modes have been studied in nc-Si superlattices with different thicknesses of nc-Si layers. Figure 2b compares the low-frequency Raman spectra in samples with 85-Å, 42-Å and 25-Å thick nc-Si layers. The spectra exhibit the expected shift toward larger wavenumbers^{14,15}, but also a significant decrease in the Raman peak-to-valley ratio. This result indicates a qualitative difference in the scattering from FLA modes for samples with different Si nanocrystal sizes. This difference may be associated with (1) an increase of the nc-Si/SiO₂ interface roughness, and/or (2) changes in the average Si nanocrystals

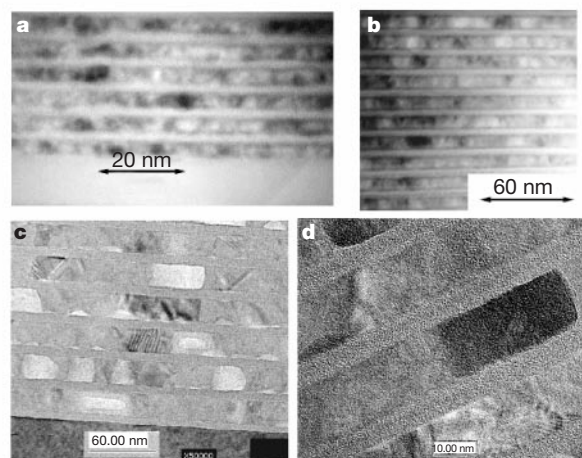


Figure 1 Transmission electron microscope micrographs showing examples of nc-Si superlattices with different thicknesses of nc-Si layers: **a**, 42 Å; **b**, 85 Å; **c**, 200 Å; and **d**, the enhanced image of a brick-shaped Si nanocrystal with 200 Å and 500 Å vertical and lateral dimensions respectively. Note that crystallization is performed under identical conditions for all samples.

crystallographic orientation. For randomly oriented Si nanocrystals the averaging of sound velocities along different Si crystallographic directions (varying from $8.5 \times 10^5 \text{ cm s}^{-1}$ to $9.2 \times 10^5 \text{ cm s}^{-1}$) should broaden the FLA signal and decrease the peak-to-valley ratio. This result also indicates that Si nanocrystals with a vertical size larger than 80 Å may have a preferred crystallographic orientation. Unfortunately, samples with a nc-Si layer thicker than 100 Å are not suitable for low-frequency scattering spectroscopy because of the

very low Raman shift of the first FLA mode ($\leq 10 \text{ cm}^{-1}$).

Samples with Si nanocrystals of vertical size 200 Å were studied using polarized Raman spectroscopy in the frequency range of the optical phonons. Raman scattering is determined by Raman polarizability tensors (R_i). Therefore, the Raman intensity depends on the orientations of polarization vectors of the incident and scattered light relative to crystallographic axes of the sample¹⁶. The crystallographic orientation of Si nanocrystals can be accurately determined by examining the anisotropy of the Raman scattering intensity. Polarized Raman experiments can be performed in an exact backscattering geometry, with wavevectors of the incident light (k_i) and scattered light (k_s) normal to the sample surface. The Raman scattering intensity in an octahedral symmetry crystal is given by¹⁶

$$S = \eta \sum_j (\mathbf{e}_i R_j \mathbf{e}_s)^2 \quad (1)$$

where η is a constant of proportionality related to the standard Raman cross-section, which depends on the frequency of the scattered light and the intensity of the incident laser; $\mathbf{e}_i$ and $\mathbf{e}_s$ are unit polarization vectors of the incident and scattered light, respectively. Following the theory of polarized Raman scattering¹⁶, the determination of the crystallographic orientation of the sample requires the measurement of the Raman scattering intensity as a function of the angle between the polarization directions of the incident and scattered light.

Figure 3 compares the measured Raman intensity as a function of the rotation angle of a polarizer (dots) for a 60-period nc-Si superlattice with a vertical nanocrystal size of 200 Å and single-crystal Si with $\langle 111 \rangle$ orientation. The experimental data (dots) were fitted (solid lines) and crystallographic orientations were calculated using analytical expressions developed in ref. 16. These measurements and calculations (Fig. 3) prove that the majority of Si nanocrystals in this superlattice have a preferred $\langle 111 \rangle$ crystallographic orientation in the direction of superlattice growth (that is, vertical direction) with less than 5% deviation.

These experiments show that the solid-phase crystallization of a nanometre-thick a-Si layer sandwiched between SiO_2 layers is accompanied by lateral self-organization, which leads to the formation of Si nanocrystals with well-defined shapes, lateral dimensions and crystallographic orientation. In part, the lateral self-organization can be expected from the constrained Si nanocrystal vertical growth, as well as by the significant difference in the rate of crystallization in different crystallographic directions¹⁷. However, it has long been held that the shape of Si nanocrystals embedded in amorphous SiO_2 should be nearly spherical owing to competition between surface and volume tensions^{6–10}. Our results

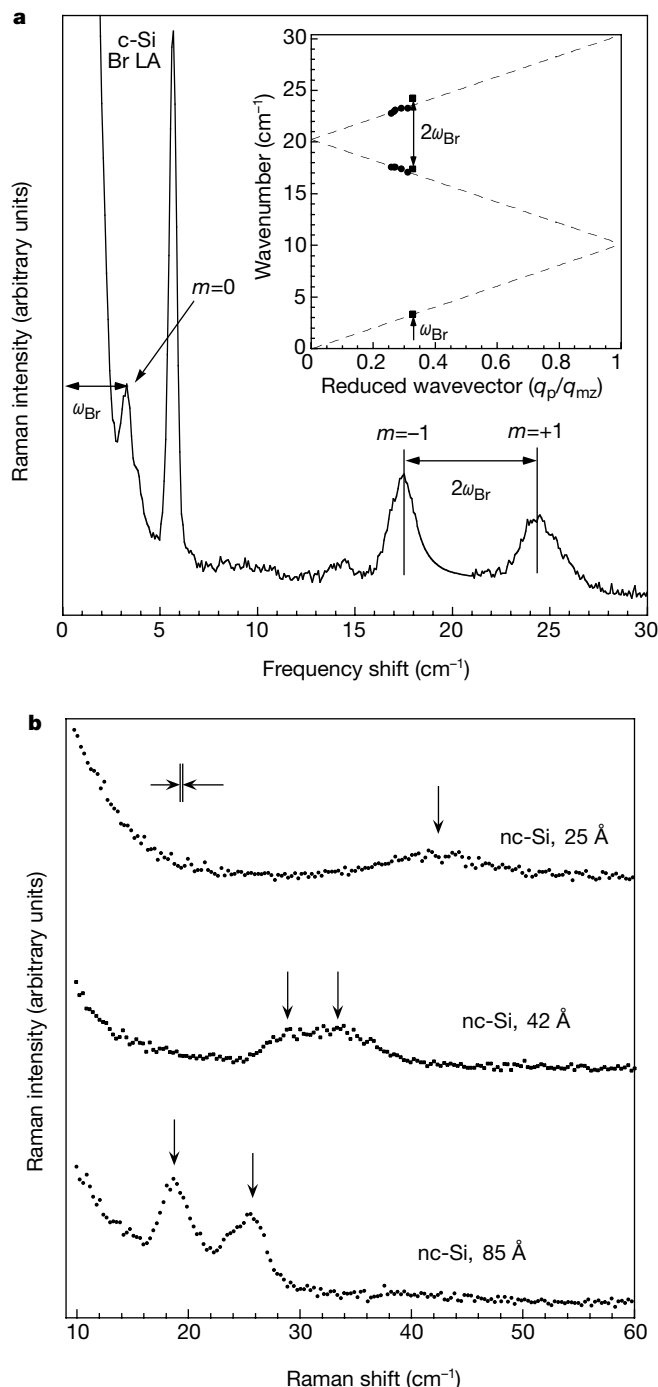


Figure 2 Spectra of scattered light. **a**, A high-resolution ($<0.06 \text{ cm}^{-1}$) spectrum of inelastically scattered light in a 20-period 85-Å nc-Si/35-Å a-SiO₂ superlattice shows Brillouin (Br) scattering from the superlattice ($m=0$) and crystalline Si substrate, and Raman scattering from the first pair of FLA phonons ($m=\pm 1$). The partial reconstruction of the Brillouin minizone (mz) using the excitation at different wavelengths is shown in the inset. **b**, Moderate resolution ($\sim 0.1 \text{ cm}^{-1}$) spectra show the Raman scattering from the first doublet of folded longitudinal acoustic phonons for samples with different nc-Si layer thickness.

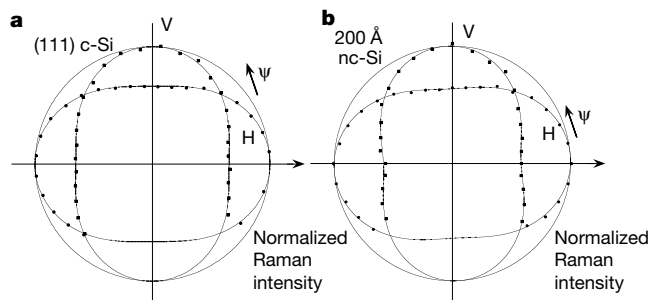


Figure 3 The Raman scattering polarization diagram. (that is, the integrated intensity of polarized Raman scattering as a function of the angle between polarizations of the incident and scattered light) in a $\langle 111 \rangle$ oriented single crystal Si (**a**) and a nc-Si superlattice with 200-Å thick nc-Si layers (**b**). The polarization of the scattered light was fixed at being either parallel or perpendicular to the entrance slit of the spectrometer, that is, H (horizontal) and V (vertical) configurations, respectively.

show that this is not the case for the solid-phase crystallization of nanometre-thick a-Si layers sandwiched between SiO₂ layers. The immediate importance of the self-organization, resulting in Si nanocrystal distinct shape, lateral size and crystallographic orientation, can be demonstrated by comparative studies of photoluminescence (PL) in isotropic Si nanoparticles (that is, Si nanocrystals smaller than 60 Å) and brick-shaped Si nanocrystals with the preferred (111) vertical crystallographic orientation.

Figure 4 compares low-temperature PL spectra from c-Si and from nc-Si superlattices with nanocrystal vertical sizes ranging from 42 Å to 200 Å. Samples with 85 Å and 200 Å Si nanocrystals exhibit PL spectra with approximately the same peak position, but the PL peak from 85 Å Si is at least twice the width compared to the 200-Å thick nc-Si layer. The PL spectrum from about 60-Å Si nanocrystals is strongly blueshifted (up to approximately 240 meV) with a full-width at half-maximum (FWHM) of the order of 80 meV. The PL spectrum in 42-Å Si nanocrystals has a peak near 1.45 eV with the FWHM of about 300 meV. To prove that this broad PL with the peak at 1.45 eV originates within Si nanocrystals, resonantly excited PL spectra were taken using tunable Ti:Al₂O₃ laser excitation. (The resonantly excited PL spectroscopy is shown to be a universal tool to study the PL origin in a system containing Si nanocrystals. For details, see ref. 18.) A clearly observable step-like structure of the PL spectrum exhibits the energies of characteristic phonons in c-Si, ruling out a defect-related PL origin¹⁸. We think our results indicate a correlation between the broadening of PL spectra and the deviation in Si nanocrystal shape, lateral size and crystallographic orientation. In addition, our data resolve an old controversy: in Si nanocrystals prepared by the high-temperature aerosol technique with a size distribution of better than 10% the PL spectrum has a FWHM of about 350 meV (refs 8 and 9), but in porous Si samples with a nanocrystal size distribution of 50–100% the typical PL FWHM is about 250–300 meV (refs 18, 19). Using our data we can immediately point out that the PL in porous Si is probably coming

from the Si single-crystal skeleton^{19,20} remaining after the electrochemical etching. In contrast, the crystallographic orientation of synthetic Si nanocrystals is believed to be random^{8,9}, thereby producing the broader PL.

We would like to emphasize that silicon-based microfabrication has been utilizing unique properties of the Si/SiO₂ interface since the beginning of complementary metal-oxide-silicon (CMOS) technology. The key issue for the development of reliable nanoscale Si structures and devices is the quality of the interfaces between Si nanocrystals and the SiO₂ insulating layer. For many applications, including Si/SiO₂ tunnel devices, a defect-free, atomically flat and compositionally abrupt Si/SiO₂ interface is absolutely essential. The creation of Si nanocrystals with a preferred crystallographic orientation, well-defined shape and crystal facets is crucial for a better understanding of these nanoscale object properties and for future developments in the area of Si/SiO₂ nanoscale devices. □

Methods

The a-Si/a-SiO₂ superlattices were grown on Si wafers by radio-frequency magnetron sputtering and plasma oxidation in a Perkin-Elmer 2400 sputtering system. The layer thickness for both a-Si and a-SiO₂ was kept constant throughout the superlattice. The a-Si layers were crystallized by rapid thermal annealing at 900 °C for 60 s followed by quasi-equilibrium furnace annealing starting from 750 °C with a 10 °C min⁻¹ temperature ramp-up to 1,100 °C. Rapid thermal and furnace annealing were performed in a nitrogen atmosphere.

The low-frequency Raman spectra were measured at room temperature in a 90° scattering geometry with the sample surface inclined at an angle of 12.3° to the incident light (which corresponds to Brewster's angle in c-Si for 457.9 nm excitation light). The large refractive index of the superlattice means that this configuration is close to the exact backscattering inside the sample. The incoming and scattered lights were polarized in the scattering plane. The excitation source was 300 mW of 457.9 nm argon-ion laser light, and the scattered light was analysed with a SOPRA DMDP2000 double monochromator and a Spex 14018 double spectrometer with a cooled RCA 34031A photomultiplier. All measurements were carried out at room temperature in a helium atmosphere.

Polarized Raman measurements were performed in the exact backscattering geometry by varying the polarization of the incident light with a step of 10°. The polarization of the scattered light was fixed at being either parallel or perpendicular to the entrance slit of the spectrometer, that is, H (horizontal) and V (vertical) configurations, respectively. For each polarization configuration the spectrum in the 505–535 cm⁻¹ range was measured with a 0.5 cm⁻¹ increment, and then numerically integrated. The spectral slit width of the spectrometer was less than 1 cm⁻¹.

The PL spectra were taken with an Acton Research Spectrometer using a North-Coast Ge detector or CCD camera. Both detectors were cooled to 77 K. The excitation sources were the 57.9 nm Ar⁺ laser line or a tunable Ti:Al₂O₃ laser. All PL measurements were done at low temperatures (4.2–15 K).

Received 25 May; accepted 17 July 2000.

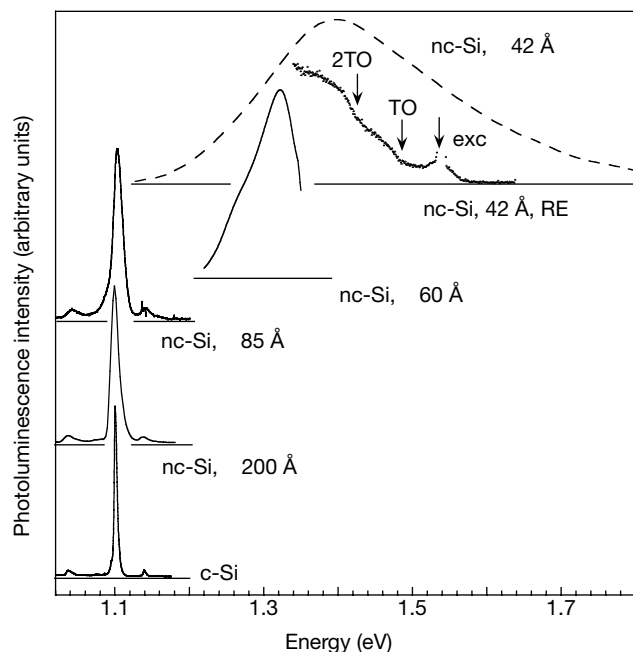


Figure 4 Photoluminescence (PL) spectra in c-Si and nc-Si superlattices with different nc-Si layer thicknesses. All non-resonantly excited PL spectra (solid and dashed lines) were recorded under similar conditions using 457.9 nm Ar⁺ laser excitation. The resonantly excited PL spectrum (dots) was taken with about 1.54-eV excitation using a tunable Ti:Al₂O₃ laser. The resonant excitation energy and the energies of optical phonons in c-Si are shown by arrows. The spectra baselines are indicated and shifted for clarity. TO, transverse optical phonon; exc, excitation.

1. Yip, S. Nanocrystals—the strongest size. *Nature* **391**, 532–533 (1998).
2. Peng, X. G. et al. Shape control of CdSe nanocrystals. *Nature* **404**, 59–61 (2000).
3. Empedocles, S. A., Neuhauser, R. & Bawendi, M. G. Three-dimensional orientation measurements of symmetric single chromophores using polarization microscopy. *Nature* **399**, 126–130 (1999).
4. Han, W. Q., Fan, S. S., Li, Q. Q. & Hu, Y. D. Synthesis of gallium nitride nanorods through a carbon nanotube-confined reaction. *Science* **277**, 1287–1289 (1997).
5. Budai, J. D. et al. Controlling the size, structure and orientation of semiconductor nanocrystals using metastable phase recrystallization. *Nature* **390**, 384–386 (1997).
6. Baron, T., Martin, F., Mur, P., Wyon, C. & Dupuy, M. Silicon quantum dot nucleation on Si₃N₄, SiO₂ and SiO₂N_x substrates for nanoelectronic devices. *J. Cryst. Growth* **209**, 1004–1008 (2000).
7. Nakajima, A., Sugita, Y., Kawamura, K., Tomita, H. & Yokoyama, N. Microstructure and optical absorption properties of Si nanocrystals fabricated with low-pressure chemical-vapor deposition. *J. Appl. Phys.* **80**, 4006–4011 (1996).
8. Wilson, W. L., Szajowski, P. F. & Brus, L. E. Quantum confinement in size-selected, surface oxidized silicon nanocrystals. *Science* **262**, 1242–1244 (1993).
9. Littau, K. A., Szajowski, P. F., Muller, A. J., Kortan, A. R. & Brus, L. E. A luminescent silicon nanocrystal colloid via a high-temperature aerosol reaction. *J. Phys. Chem.* **97**, 1224–1230 (1993).
10. Shimizu-Iwayama, T., Hole, D. E. & Boyd, I. W. Mechanism of photoluminescence of Si nanocrystals in SiO₂ fabricated by ion implantation: the role of interactions of nanocrystals and oxygen. *J. Phys. Condens. Matter* **11**, 6595–6604 (1999).
11. Tsybeskov, L. et al. Nanocrystalline-silicon superlattice produced by controlled recrystallization. *Appl. Phys. Lett.* **72**, 43–45 (1998).
12. Levi, D., Zhang, S.-L., Klein, M. V., Klem, J. & Morkoc, H. Raman study of the effects of annealing on folded LA and confined LO phonons in GaAs-AlAs superlattices. *Phys. Rev. B* **36**, 8032–8037 (1987).
13. Lockwood, D. J., Dharma-wardana, M. W. C., Baribeau, J.-M. & Houghton, D. C. Folded acoustic phonons in Si/Ge_{1-x} strained superlattices. *Phys. Rev. B* **35**, 2243–2251 (1987).
14. Ivchenko, E. I. & Pikus, G. *Superlattices and Other Heterostructures: Symmetry and Optical Phenomena* Vol. 110 (Springer Ser. Solid-State Science, Springer, Berlin, 1995).
15. Ghanbari, R. A., White, J. D., Fasol, G., Gibbings, C. J. & Tuppen, C. G. Phonon frequency for Si-Ge strained-layer superlattices calculated in a 3-dimensional model. *Phys. Rev. B* **42**, 7033–7041 (1990).

16. Mizoguchi, K. & Nakashima, S. J. Determination of crystallographic orientations in silicon films by Raman-microprobe polarization measurements. *J. Appl. Phys.* **65**, 2583–2590 (1989).
17. Haji, L., Joubert, P., Stoemenos, J. & Economou, N. A. Mode of growth and microstructure of polycrystalline silicon obtained by solid phase crystallization of an amorphous silicon film. *J. Appl. Phys.* **75**, 3944–3952 (1994).
18. Calcott, P. D. J., Nash, K. J., Canham, L. T., Kane, M. J. & Brumhead, D. Spectroscopic identification of the luminescence mechanism of highly porous silicon. *J. Lumin.* **57**, 257–269 (1993).
19. Cullis, A. G., Canham, L. T. & Calcott, P. D. The structural and luminescence properties of porous silicon. *J. Appl. Phys.* **82**, 909–965 (1997).
20. Cullis, A. G. & Canham, L. T. Visible light emission due to quantum size effects in highly porous crystalline silicon. *Nature* **353**, 335–338 (1991).

Acknowledgements

This work is supported by US Army Research Office, National Science Foundation and Motorola. L.T. acknowledges support from the German Academic Exchange Service (DAAD).

Correspondence and requests for materials should be addressed to L.T.
(e-mail: tsybeskov@ece.rochester.edu).

Direct electrochemical reduction of titanium dioxide to titanium in molten calcium chloride

George Zheng Chen*, Derek J. Fray* & Tom W. Farthing†

* Department of Materials Science and Metallurgy, University of Cambridge, Pembroke Street, Cambridge CB2 3QZ, UK

† 71 Sir Richards Drive, Harborne, Birmingham B17 8SG, UK

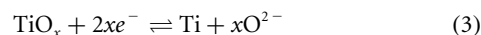
Many reactive metals are difficult to prepare in pure form without complicated and expensive procedures^{1–18}. Although titanium has many desirable properties (it is light, strong and corrosion-resistant¹), its use has been restricted because of its high processing cost. In the current pyrometallurgical process—the Kroll process^{4,5}—the titanium minerals rutile and ilmenite are carbochlorinated to remove oxygen, iron and other impurities, producing a TiCl₄ vapour. This is then reduced to titanium metal by magnesium metal; the by-product MgCl₂ is removed by vacuum distillation. The prediction that this process would be replaced by an electrochemical route^{6–10} has not been fulfilled; attempts involving the electro-deposition of titanium from ionic solutions have been hampered by difficulties in eliminating the redox cycling of multivalent titanium ions and in handling very reactive dendritic products^{4,5}. Here we report an electrochemical method for the direct reduction of solid TiO₂, in which the oxygen is ionized, dissolved in a molten salt and discharged at the anode, leaving pure titanium at the cathode. The simplicity and rapidity of this process compared to conventional routes should result in reduced production costs and the approach should be applicable to a wide range of metal oxides.

The electrochemical deoxygenation of thin titanium foils in molten CaCl₂ has been described^{19,20}. The proposed mechanism is that when calcium is deposited on the titanium foil cathode, it reacts with the oxygen in the foil to form CaO, which is soluble in molten CaCl₂ (ref. 21). An alternative explanation is that oxygen ionization occurs at a less cathodic (negative) potential than calcium deposition and so direct reduction of titanium oxides to titanium metal can be achieved electrochemically rather than by the chemical reaction with calcium. A comparison between the values of Gibbs free energy of stoichiometric titanium oxides and CaO agrees with this hypothesis. These two concepts of deoxygenation may be summarized as follows.

(1) Deposition of calcium at a more cathodic potential, E_1^0



(2) Ionization of oxygen at a less cathodic potential, E_2^0



In order to test the feasibility of the concept of oxygen ionization in equation (3), several experiments were performed. Firstly, oxide scales (5–50 μm) were formed on the surfaces of chemically pure titanium foils at 700–900 °C in air. These oxide scales consisted of mainly TiO₂ as confirmed by scanning electron microscopy (SEM) and energy dispersive X-ray (EDX) analysis. Cyclic voltammetry was then performed in molten CaCl₂ to find the ionization potential of oxygen in the scale. Figure 1 compares the cyclic voltammograms of two titanium foils with or without the scale. Because of the voltage loss across the oxide scale, the potentials of the cathodic current waves in Fig. 1a are shifted negatively in comparison with those in Fig. 1b. Similarly, the anodic waves are shifted positively in Fig. 1a than in Fig. 1b.

Waves 4, 4', 8 and 8' appear at potentials beyond which the current increases continuously. This phenomenon is typically the same as those resulting from the decomposition of a solvent or electrolyte and/or dissolution of electrode, which defines the 'potential window' in cyclic voltammetry. Therefore, we can confidently attribute waves 4 and 4' to the deposition of calcium metal, and waves 8 and 8' to the dissolution of titanium.

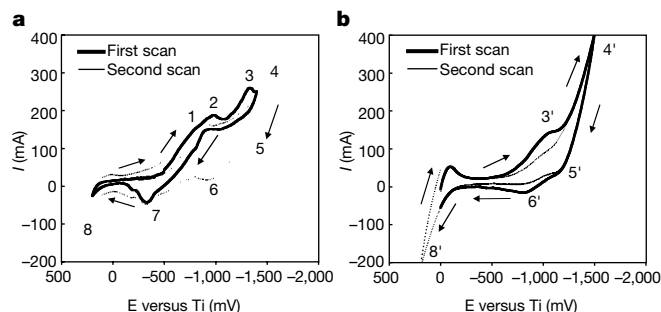


Figure 1 Cyclic voltammograms of titanium foils in molten CaCl₂. **a**, Oxide-scale-coated titanium foil, prepared by heating in air at 700 °C for 14 days; electrolyte-submerged surface area: ~1.0 cm². **b**, The 'as received' titanium foil which was used to prepare the scale-coated electrode in **a** with electrolyte-submerged surface area: ~1.5 cm². Both voltammograms were obtained at 10 mV s⁻¹ and 800 °C under argon. The arrows show the directions of current and potential variation when the potential is scanned forward (negative direction) and backward (positive direction).

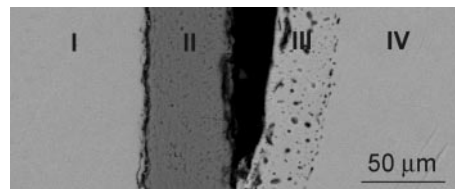


Figure 2 Reduction of titanium oxide scale to titanium metal. Back-scattered scanning electron microscopy (SEM) images of the cross-section of a scale-coated titanium foil (700 °C, 14 days) before (left) and after (right) electrolysis in molten CaCl₂ (3.0 V, 950 °C, 4 hours). **I** and **IV**: bulk metal regions; **II**: oxide scale; **III**: porous metal layer. Energy dispersive X-ray (EDX) and microhardness analyses confirm that the oxygen concentration in **I** is increasing from the left end (Vickers hardness number VHN = 200–240) to the interface between **I** and **II** (VHN > 800); in **II** agrees with the stoichiometry of titanium oxides; in **III** and **IV** is the same as or lower than that in the 'as received' titanium foil (VHN = 200–240).

Current waves 1 and 2, and their likely anodic counterpart wave 7 clearly originate from the oxide scale as the three waves are absent on the bare titanium foil. To determine which wave is responsible for reducing the oxide scale, chronoamperometry was performed with the electrode potential being set in the region of wave 2. In the experiment, the temperature was adjusted to 950 °C to increase the reaction rate. The working electrode potential was accordingly lowered to -0.85 V as indicated by the voltammograms recorded at 950 °C. The electrolysis lasted for about an hour before removing the working electrode from the electrolyte. After cooling in air and washing with water, the electrode exhibited a metallic appearance on the surface and optical microscopy on the cross-section of the foil further confirmed that the oxide layer was transformed into a porous metal layer. Microhardness tests, however, revealed that the oxygen level in the metal region next to the porous surface layer had increased significantly, indicating the diffusion of oxygen from the oxide scale into the neighbouring metal phase. These results clearly demonstrate that at potentials in the region of wave 2, only the oxygen from titanium oxides can be ionized, but to remove the oxygen dissolved in the metal requires more negative potentials. This finding implies that the coupled waves 3 and 6 (Fig. 1a) and waves 3' and 6' (Fig. 1b) are related to one of the oxygenated metal phases as indicated by the oxygen-titanium phase diagram³. Further, we found that waves 3' and 6' were repeatable by heating titanium in an oxygen-deficient environment where oxygen dissolution into the metal occurred but no oxide formed. By extending the cathodic limit of the potential scan, the second scan on the scale-coated foil shows a great increase in the anodic wave 5 after the scan

is reversed. This wave is the same as wave 5' in Fig. 1b, obviously due to the re-dissolution of the deposited calcium metal. By comparing the first and second scans on the scale-coated foil, it is certain that when the electrode potential is not more negative than wave 3, no calcium metal will form, but oxygen can be removed using only the ionization mechanism.

Following the voltammetric findings, the scale-coated titanium foils were made the cathode and electrolysed with a graphite rod anode in molten CaCl₂ in a titanium crucible at 950 °C. The surface area of the cathode in contact with electrolyte was about 3.0 cm². The voltage applied was 2.8–3.2 V, well beneath the decomposition voltage of CaCl₂ considering the voltage losses due to the resistance of electrodes, electrolyte and other materials used in the electric circuit (the 'IR drop'). The responding current depended on the applied voltage, cathode surface area and scale thickness, but was not too far away from 1.0 A in the beginning and dropped gradually to the background value, ~0.2 A. After a time between 30 min to a few hours, the electrolysis was terminated and the cathode removed from the melt at a temperature just above the melting point of CaCl₂. After washing in water, it was seen that the surface of the cathode exhibited a metallic appearance. The cross-section of cathode was then investigated using microhardness testing, SEM and EDX. An example is displayed in Fig. 2, showing that the oxide scale has completely changed into a porous metal layer, accompanied by shrinkage. The electrolysed samples were also analysed by vacuum fusion elemental analysis with the detected oxygen concentration varying from that in the 'as received' titanium foils down to below the detection limit (200 p.p.m.).

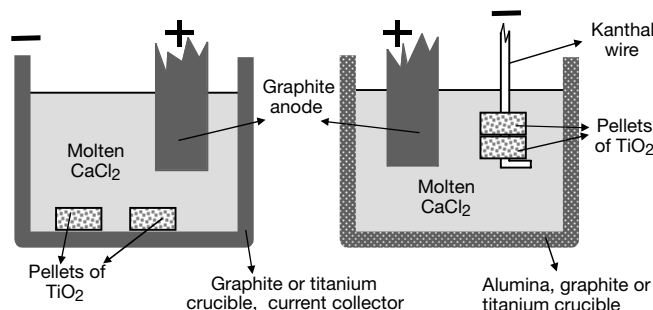


Figure 3 Electrolytic cells for the reduction of TiO₂ pellets. The left cell shows two pellets placed at the bottom of a titanium or graphite crucible which also functions as the cathodic current collector. This arrangement resembles a basket electrode that can be removed from the cell after electrolysis, allowing filtration of molten salt. The two pellets in the right

cell are threaded onto a Kanthal wire (current collector) and suspended in the molten salt. This type of assembled electrode can also be removed from the cell easily after electrolysis.

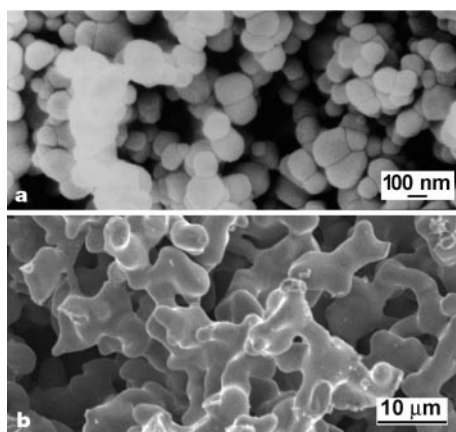


Figure 4 Reduction of the TiO₂ powder to titanium metal. SEM images of the microstructures in a pellet of pressed (750 kg cm⁻², slow pressure release) and sintered (850 °C, 48 hours) TiO₂ powder (+99.9% purity) before (a) and after (b) placing it in a titanium crucible filled with molten CaCl₂ and performing electrolysis (3.2 V, 950 °C,

12 hours) under argon with a graphite rod anode. EDX analysis revealed the stoichiometric composition of TiO₂ in a, but that the oxygen concentration in b fell below the detection limit.

These experiments indicated that it might be possible to electrochemically reduce TiO_2 directly to the metal in molten CaCl_2 . To confirm this, small TiO_2 pellets of 5–10 mm diameter and 2–10 mm thickness were made from the powder (see the Methods section for preparation details) and placed or suspended on a Kanthal wire in a titanium, graphite or alumina (for suspended pellets only) crucible filled with molten CaCl_2 at 850–950 °C (see Fig. 3). A voltage of 3.0–3.2 V was applied between a graphite rod anode and the wall of the crucible or the Kanthal wire, resulting in a high initial current of $\sim 10^4 \text{ A m}^{-2}$ calculated from the geometric surface area of the pellets. This current gradually decayed to a limiting value (independent of the numbers of pellets) over a period of 5–24 hours depending on the size and numbers of pellets. For relatively larger pellets suspended on the Kanthal wire, it was observed that the current momentarily increased and then decreased slowly. After electrolysis and washing, the pellets exhibited a metallic grey colour. Examinations under SEM showed that the 0.25- μm oxide powder (see Fig. 4a) had been transformed into titanium metal particles of about 12 μm which were slightly sintered together to give the familiar microstructure of the Kroll titanium sponge (Fig. 4b).

It is surprising that TiO_2 , which is effectively an insulator, can be reduced so quickly. However, examination of the phase diagram³ shows that the removal of a small amount of oxygen from TiO_2 results in the formation of Magnelli phases (TiO_{2-x}) which are highly conducting²². It seems that upon applying the voltage, a small amount of reduction takes place in the porous pellet, causing it to become an electronically conducting electrode. This explains the observed initial increase in current when electrolysis larger pellets. The ionization of oxygen then takes place throughout the pellet, with the oxygen anions dissolving in CaCl_2 , where they diffuse quickly out of the pellet into the melt and eventually discharge at the anode. The rate-limiting step is likely to be the diffusion of the oxygen in the solid phases.

We have thus investigated a new method of reducing solid TiO_2 to titanium in which the TiO_2 powder is made the cathode in molten CaCl_2 , whose cation can form a more stable oxide, CaO . The oxygen in TiO_2 is simply ionized and dissolves in the salt, leaving titanium metal behind. This route seems very much easier and quicker than the established routes for many metals. This is only a laboratory study but our observations have been repeated by an independent laboratory on a kilogram scale, proving that it is possible to scale up this technology (A. Godfrey and C. M. Ward-Close, personal communication).

It is our anticipation that this technique can be applied to other metal oxides. Individual or mixed oxide powders of Al, B, Cr, Fe, Nd and V have already been reduced (G.Z.C. and D.J.F., unpublished results). The work is being extended to the oxides of other elements of groups III–X and rare earths. □

Methods

Grey-blue to earth-yellow oxide scales were formed by heating free-standing chemically pure titanium foils (50–80 mm length, 4–5 mm width, 1–2 mm thickness, 2,000–3,000 p.p.m. oxygen) in air at 700–900 °C for 1–340 hours. Coherent scales were always obtained when the temperature was not higher than 700 °C. Beyond 800 °C, a coherent scale can only be obtained if the heating period was short. The scales were analysed by SEM and EDX, confirming that the composition was mainly TiO_2 , in agreement with ref. 23.

Electrode preforms of TiO_2 suitable for molten salt electrolysis were prepared from TiO_2 powders (various grades of both anatase and rutile, particle size: 0.25–5 μm) by pressing or slip-casting. The results from pressing and slip casting were not significantly different, although slip casting was preferred in this work as it was quick, suitable for making preforms of various shapes and sizes, and also convenient when mixing different powders. After sintering at 800–950 °C in air for 2–48 hours (the higher the temperature, the shorter the time), the strength of the preforms increased, and holes could be drilled in them, allowing them to be threaded onto a current collector, such as Kanthal wire. The porosity of sintered pellets was estimated from the weight to volume ratio of the pellets and the density of rutile, and further confirmed by measuring the weight difference of the pellets in air and in pure water. Such obtained porosity values were typically 40–60%, depending on the pressure of pressing or density of the slurry and also the sintering temperature and time.

The acidic slurry of amorphous TiO_2 formed by treating ilmenite with sulphuric acid was also used for preparing the TiO_2 preforms. The sintered pellets showed significant shrinkage and greatly increased strength in comparison with those prepared from the powders.

All molten salt experiments were performed in a sealed Inconel reactor (10 cm inner diameter, 57 cm height) which was continuously purged with argon ($150\text{--}180 \text{ ml min}^{-1}$). Pre-electrolysis for removing moisture from the molten salt was carried out at 2.5–2.7 V for 2 hours or longer, between the graphite anode and the crucible or a titanium foil or Kanthal wire cathode.

Cyclic voltammetry and chronoamperometry (controlled potential electrolysis) were performed in the three-electrode manner in molten CaCl_2 with an oxide-scale-coated titanium foil (1.0–1.5 cm^2 in electrolyte) as the working electrode and a graphite rod (1 cm diameter, 3 cm depth in electrolyte) as the counter electrode. A titanium crucible (10 cm depth, 5 cm inner diameter, 0.5 cm wall thickness) was used as the container for the molten salt and also functioned as a pseudo-reference electrode whose potential was often checked against the potentials for calcium deposition and titanium dissolution. The checking was carried out on an independent titanium foil (99.7%, Aldrich). It was observed that the potential of the pseudo-reference was satisfactorily stable over a period of up to 2 hours.

The electrolysis of scale-coated titanium foils and TiO_2 pellets were performed in the two-electrode manner in molten CaCl_2 contained in a crucible of titanium, alumina or graphite (10–15 cm depth, 2.5–8 cm inner diameter). The quantities of TiO_2 pellets and molten salt were usually 1–3 g and 150–200 g respectively. The anode gases of the electrolysis included oxygen and carbon oxides as determined by sensors and gas chromatography. This is in accordance with other studies of carbon with oxygen at moderate temperatures²⁴. After electrolysis, the titanium foils or suspended pellets were removed from the crucible at a temperature just above the melting point of the melt, allowed to cool, and then washed with water. For pellets placed at the bottom of a conducting crucible, after electrolysis, they were cooled together with the melt to room temperature and then the solidified salt was extracted into water by stirring and the pellets further washed.

Voltagages (2.8–3.2 V) for electrolysis were chosen such that, taking the IR drop into consideration, the decomposition of CaCl_2 (3.2–3.3 V) would not occur but the reduction of all likely involved oxides could take place. The total circuit and electrolyte resistance of our cell was about 0.5–1.0 ohm, as measured by the a.c. impedance method.

Received 31 January; accepted 31 July 2000.

- Habashi, F. (ed.) *Handbook of Extractive Metallurgy* 1129–1180 (Wiley-VCH, Weinham, 1997).
- Sadoway, D. R. in *Refractory Metals Extraction, Processing and Applications: Proc. TMS Annual Meeting, New Orleans 1991* (eds Liddel, K. C., Sadoway, D. R. & Bautista, R. G.) (TMS—The Minerals, Metals & Materials Society, Warrendale, 1991).
- Massalski, T. B., Okamoto, H., Subramanian, P. R. & Kacprzak, L. (eds) *Binary Alloy Phase Diagrams* 2nd edn, Vol. 3, 2924–2927 (ASM International, Materials Park, 1990).
- Kroll, W. J. The production of ductile titanium. *Trans. Am. Electrochem. Soc.* **78**, 35–47 (1940).
- Ikeshima, T., in *Titanium Science and Technology, Proc. 5th Int. Conf. Titanium, München 1984* (eds Lufjering, G., Zwicker, U. & Bunk, W.) 3–14 (DGM-Deutsche Gesellschaft für Materialkunde e.V., Oberursel, 1985).
- Cobel, G., Fisher, J. & Synder, L. E. in *Titanium '80, Science and Technology, Proc. 4th Int. Conf. Titanium, Kyoto 1980* (eds Kimura, H. & Izumi, O.) 1969–1976 (The Metallurgical Society of AIME, Warrendale, 1980).
- Opie, W. R. & Moles, O. W. A basket cathode electrolytic cell for production of titanium. *Trans. Met. Soc. AIME* **218**, 646–649 (1960).
- Ginatta, M. V. Method of producing metals by cathodic dissolution of their compounds. US Patent 4,400,247 (23 Aug. 1983).
- Froes, F. H. Titanium and other light metals: let's do something about cost. *JOM* **50**, 15 (1998).
- Hartman, A. D., Gerdemann, S. J. & Hansen, J. S. Producing lower-cost titanium for automotive applications. *JOM* **50**, 16–19 (1998).
- Suzuki, K. The high-quality precision casting of titanium alloys. *JOM* **50**, 20–23 (1998).
- Okabe, T., Ohkubo, C., Watanabe, I., Okuno, O. & Takada, Y. The present status of dental titanium casting. *JOM* **50**, 24–29 (1998).
- Froes, F. H. The production of low-cost titanium powders. *JOM* **50**, 41–43 (1998).
- Tapphorn, R. M. & Gabel, H. The solid-state spray forming of low-oxide titanium components. *JOM* **50**, 45–46, 76 (1998).
- Elliott, G. R. B. The continuous production of titanium powder using circulating molten salt. *JOM* **50**, 48–49 (1998).
- Sohn, H. Y. Ti and TiAl powders by the flash reduction of chloride vapors. *JOM* **50**, 50–51 (1998).
- Segall, A. E., Papyrin, A. N., Conway, J. C. Jr & Shapiro, D. A cold-gas spray coating process for enhancing titanium. *JOM* **50**, 52–54 (1998).
- Larson, H. R. & Eagar, T. W. The plasma-enhanced recovery of titanium by the electrolysis of titanate slags. *JOM* **50**, 56–57 (1998).
- Okabe, T. H., Oishi, T. & Ono, K. Deoxidation of titanium aluminide by Ca–Al alloy under controlled aluminum activity. *Metall. Trans. B* **23**, 583–590 (1992).
- Okabe, T. H., Deura, T., Oishi, T., Ono, K. & Sadoway, D. R. Thermodynamic properties of oxygen in yttrium–oxygen solid solutions. *J. Alloys Compounds* **237**, 841–847 (1996).
- Boghosian, S., Goto, A., Mediaas, H., Ravlo, W. & Ostvold, T. Oxide complexes in alkali alkaline-earth chloride melts. *Acta Chem. Scand.* **45**, 145–157 (1991).
- Gruber, H. & Krautz, E. Magnetoresistance and conductivity in the binary-system titanium oxygen. 2. Semiconductive titanium-oxides. *Phys. Status Solidi A* **69**, 287–295 (1982).
- McQuillan, A. D. & McQuillan, M. K. *Titanium* 402–426 (Butterworths Scientific, London, 1956).
- McKee, D. W. Gasification of graphite in carbon-dioxide and water-vapour—the catalytic effects of alkali-metal salts. *Carbon* **20**, 59–66 (1982).

Acknowledgements

This research was sponsored by the EPSRC. T.W.F. first suggested the electrochemical deoxidation of titanium metal. G.Z.C. was the first to observe that it was possible to reduce thick layers of oxide on titanium metal using molten salt electrochemistry. D.J.F. suggested the experiment, which was carried out by G.Z.C., on the reduction of the solid titanium dioxide pellets. M. S. P. Shaffer took the original SEM image of Fig. 4a.

Correspondence and requests for materials should be addressed to D. J. F. (e-mail: djf25@hermes.cam.ac.uk).

Reduced calcification of marine plankton in response to increased atmospheric CO₂

Ulf Riebesell*, Ingrid Zondervan†, Björn Rost*, Philippe D. Tortell‡, Richard E. Zeebe*§ & François M. M. Morel†

* Alfred Wegener Institute for Polar and Marine Research, P.O. Box 120161, D-27515 Bremerhaven, Germany

† Department of Geosciences & Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544, USA

‡ Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA

The formation of calcareous skeletons by marine planktonic organisms and their subsequent sinking to depth generates a continuous rain of calcium carbonate to the deep ocean and underlying sediments¹. This is important in regulating marine carbon cycling and ocean-atmosphere CO₂ exchange². The present rise in atmospheric CO₂ levels³ causes significant changes in surface ocean pH and carbonate chemistry⁴. Such changes have been shown to slow down calcification in corals and coralline macroalgae^{5,6}, but the majority of marine calcification occurs in planktonic organisms. Here we report reduced calcite production at increased CO₂ concentrations in monospecific cultures of two dominant marine calcifying phytoplankton species, the coccolithophorids *Emiliania huxleyi* and *Gephyrocapsa oceanica*. This was accompanied by an increased proportion of malformed coccoliths and incomplete coccospheres. Diminished calcification led to a reduction in the ratio of calcite precipitation to organic matter production. Similar results were obtained in incubations of natural plankton assemblages from the north Pacific ocean when exposed to experimentally elevated CO₂ levels. We suggest that the progressive increase in atmospheric CO₂ concentrations may therefore slow down the production of calcium carbonate in the surface ocean. As the process of calcification releases CO₂ to the atmosphere, the response observed here could potentially act as a negative feedback on atmospheric CO₂ levels.

By the end of the next century, the expected increase in atmospheric CO₂ (ref. 3) will give rise to an almost threefold increase in surface ocean CO₂ concentrations relative to pre-industrial values. This will cause CO₃²⁻ concentrations and surface water pH to drop by about 50% and 0.35 units, respectively⁴. Changes of this magnitude have been shown to significantly slow down calcification of temperate and tropical corals and coralline macroalgae^{5,6}. Although coral reefs are the most conspicuous life-supporting calcareous structures, the majority of biogenic carbonate precipitation (>80%) is carried out by planktonic microorganisms¹, particularly coccolithophorids⁷. These unicellular microalgae are major contributors to marine primary production and an important component of open ocean and coastal marine ecosystems⁸. Two prominent

representatives of the coccolithophorids, *Emiliania huxleyi* and *Gephyrocapsa oceanica*, are both bloom-forming and have a world-wide distribution. *G. oceanica* is the dominant coccolithophorid in neritic environments of tropical waters⁹, whereas *E. huxleyi*, one of the most prominent producers of calcium carbonate in the world ocean¹⁰, forms extensive blooms covering large areas in temperate and subpolar latitudes^{9,11}.

The response of these two species to CO₂-related changes in seawater carbonate chemistry was examined under controlled

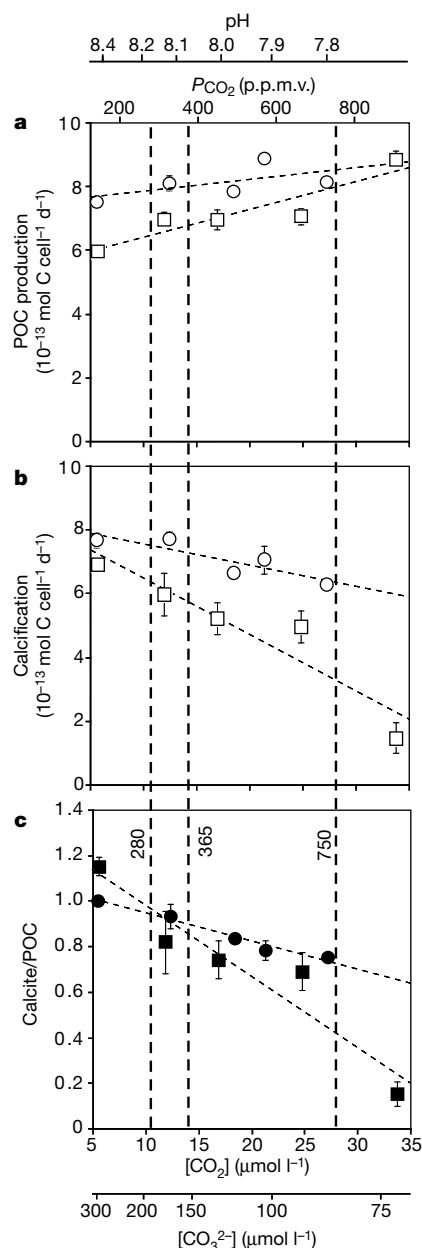


Figure 1 Response of organic and inorganic carbon production to CO₂ concentration in laboratory-cultured coccolithophorids. **a**, Particulate organic carbon (POC) production; **b**, calcification; and **c**, the ratio of calcification to POC production (calcite/POC) of the coccolithophorids *Emiliania huxleyi* (circles) and *Gephyrocapsa oceanica* (squares) as a function of CO₂ concentration, [CO₂]. Bars denote ± 1 s.d. ($n = 3$); dotted lines represent linear regressions. Also indicated are corresponding ranges of pH, p_{CO_2} , and [CO₃²⁻]. We note that DIC and total alkalinity differed slightly between experimental sets for the two species; values for pH, p_{CO_2} , and [CO₃²⁻], therefore, are only approximations. Vertical lines indicate p_{CO_2} values of 280, 365 and 750 p.p.m.v., representing pre-industrial, present day and future concentrations.

laboratory conditions. The carbonate system of the growth medium was manipulated by adding acid or base to cover a range from pre-industrial CO_2 levels (280 p.p.m.v.) to approximately triple pre-industrial values (about 750 p.p.m.v.). Over this range, *E. huxleyi* and *G. oceanica* experience a slight increase in photosynthetic carbon fixation of 8.5% and 18.6%, respectively (Fig. 1a), and a comparatively larger decrease in the rate of calcification of 15.7% and 44.7%, respectively (Fig. 1b). The ratio of calcite to organic matter production (calcite/POC) for the two species decreased by 21.0% and 52.5%, respectively, between 280 and 750 p.p.m.v. (Fig. 1c). Since calcite production has been shown to vary with ambient light conditions¹², we have grown *E. huxleyi* under different light/dark cycles and photon flux densities. A similar decrease in the calcite/POC ratio in response to CO_2 -related changes in carbonate chemistry was obtained over a fivefold range in photon flux densities (Fig. 2).

Scanning electron microscopy indicated that malformed coccoliths and incomplete coccospheres increased in relative numbers

with increasing CO_2 concentrations (Fig. 3). Coccolith undercalcification and malformation is a common phenomenon frequently observed both in natural environments and under laboratory conditions¹³. The systematic trend in the relative abundance of malformed coccoliths and coccospheres observed here, however, suggests a direct effect of seawater carbonate chemistry on the regulatory mechanisms controlling coccolith production inside the cell. Based on light microscopic analysis, no consistent trend was obtained in the number of attached or free coccoliths per coccosphere.

Our laboratory results are consistent with CO_2 -related responses of natural plankton assemblages collected in the subarctic north Pacific, a region where coccolithophorids are major contributors to primary production¹⁴. After incubation of replicate samples at p_{CO_2} levels of about 250 p.p.m.v. and about 800 p.p.m.v. for 1.5 to 9 days, the rate of calcification was reduced by 36% to 83% in high- CO_2 relative to low- CO_2 treatments in four independent experiments (Fig. 4). A similar CO_2 -dependent response was obtained under

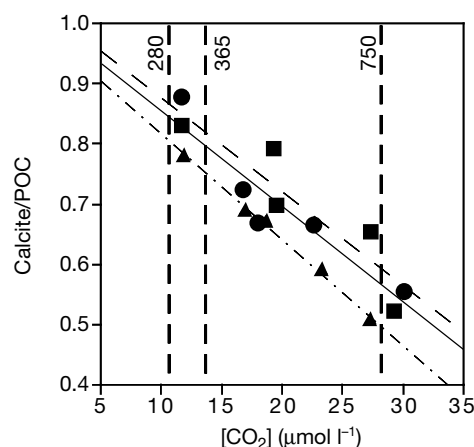


Figure 2 Ratio of calcification to POC production (calcite/POC) of *Emiliania huxleyi* as a function of CO_2 concentration, $[\text{CO}_2]$. Cells were incubated at photon flux densities of 30, 80 and $150 \mu\text{mol m}^{-2} \text{s}^{-1}$ (denoted by circles, squares and triangles and corresponding

solid, dashed, dash-dotted regression lines, respectively). Bars denote ± 1 s.d. ($n = 3$); lines represent linear regressions. Vertical lines indicate p_{CO_2} values of 280, 365 and 750 p.p.m.v.

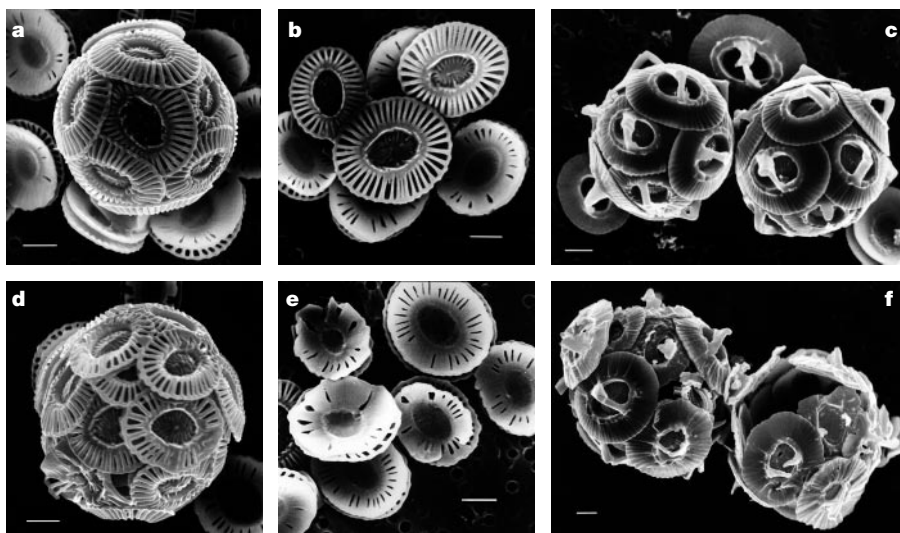


Figure 3 Scanning electron microscopy (SEM) photographs of coccolithophorids under different CO_2 concentrations. **a, b, d, e**, *Emiliania huxleyi*; and **c, f**, *Gephyrocapsa oceanica* collected from cultures incubated at $[\text{CO}_2] \approx 12 \mu\text{mol l}^{-1}$ (**a–c**) and at $[\text{CO}_2] \approx 30–33 \mu\text{mol l}^{-1}$ (**d–f**), corresponding to p_{CO_2} levels of about 300 p.p.m.v. and 780–850 p.p.m.v., respectively. Scale bars represent $1 \mu\text{m}$. Note the difference in the

coccolith structure (including distinct malformations) and in the degree of calcification of cells grown at normal and elevated CO_2 levels. Pictures are selected from a large set of SEM photographs to depict the general trend in coccolith calcification. As the culture medium was super-saturated with respect to calcium carbonate under all experimental conditions, post-formation calcite dissolution is not expected to have occurred.

reduced light intensities (10% surface irradiance, data not shown). No significant differences were obtained between short and longer-term incubations of the natural assemblages. As short-term incubations are not likely to experience large changes in species composition, the observed response most probably reflects a reduction in carbonate precipitation of the calcifying organisms in the plankton assemblage.

The observed decrease in calcification with increasing p_{CO_2} , if representative of biogenic calcification in the world's ocean, has significant implications for the marine carbon cycle. Owing to its effect on carbonate system equilibria, calcification is a source of CO_2 to the surrounding water¹⁵, whereby the increase in CO_2 concen-

tration, due to calcification is a function of the buffer capacity of sea water. Theoretically, the buffer state of pre-industrial sea water resulted in 0.63 mole CO_2 released per mole CaCO_3 precipitated¹⁶ (assuming temperature $T = 15^\circ\text{C}$, and salinity $S = 35$). Following the predictions of future atmospheric CO_2 rise, this value will increase to 0.79 in 2100 (assuming Intergovernmental Panel on Climate Change (IPCC) scenario IS92a, ref. 3). At constant global ocean calcification this results in an additional source of CO_2 to the atmosphere. In the case of reduced calcification, this positive feedback is reversed. Assuming a pre-industrial pelagic inorganic carbon production of 0.86 Gt C y^{-1} (ref. 17) and a CO_2 -related decrease in planktonic calcification as observed in our laboratory and field experiments (ranging between 16% and 83%), model calculations yield an additional storage capacity of the surface ocean for CO_2 between 6.2 Gt C and 32.3 Gt C for the period of 1950 to 2100.

Our results indicate that the ratio of calcite to organic matter production in cultured coccolithophorids and in oceanic phytoplankton assemblages is highly sensitive to the seawater p_{CO_2} . Although it is presently not clear what the physiological and ecological role of coccolith formation is¹⁸, we propose that CO_2 -dependent changes in calcification may affect cellular processes such as acquisition of inorganic carbon¹⁹ and nutrients²⁰ as well as trophic interactions, and particle sinking rate^{21,22}. These, in turn, may influence the structure and regulation of marine ecosystems in which coccolithophorids are dominant. From a geochemical viewpoint, a decrease in global ocean calcification would enhance CO_2 storage in the upper ocean^{3,15,23}, thus providing a negative feedback for changes in atmospheric p_{CO_2} . Such a feedback should be taken into account when predicting the role of the ocean in mitigating future anthropogenic CO_2 increases or in reconstructing the relation between ocean productivity and glacial–interglacial variations in p_{CO_2} .

Methods

Laboratory

Monospecific cultures of the coccolithophorids *Emiliania huxleyi* (strain PML B92/11A) and *Gephyrocapsa oceanica* (strain PC7/1) were grown in dilute batch cultures at 15°C in filtered ($0.2 \mu\text{m}$) sea water enriched with nitrate and phosphate to concentrations of 100 and $6.25 \mu\text{mol l}^{-1}$, respectively, and with metals and vitamins according to the *f/2* culture medium (ref. 24). The carbonate system was adjusted through addition of 1 N HCl or 1 N NaOH to the medium. Cells were acclimated to the experimental conditions for 7–9 generations and allowed to grow for about 8 cell divisions during experiments. Cultures were incubated in triplicate at photon flux densities of $150 \mu\text{mol m}^{-2} \text{ s}^{-1}$, light/dark (L/D) cycle = 16/8 h (Fig. 1) and of $150, 80$ and $30 \mu\text{mol m}^{-2} \text{ s}^{-1}$, L/D = 24/0 (Fig. 2). Dissolved inorganic carbon (DIC) was measured coulometrically in duplicate (UIC model 5012)²⁵. Alkalinity was determined in duplicate through potentiometric titration²⁶. pH, CO_2 and CO_3^{2-} concentrations were calculated from alkalinity, DIC and phosphate concentrations ($T = 15^\circ\text{C}$; $S = 31$) using the dissociation constants of ref. 27. Subsamples for total particulate carbon (TPC) and particulate organic carbon (POC), which in L/D = 16/8 were taken at the end of the dark phase, were filtered onto pre-combusted (12 h, 500°C) QM-A filters (pore width is about $0.6 \mu\text{m}$) and stored at -25°C . Before analysis, POC filters were fumed for 2 h with saturated HCl solution in order to remove all inorganic carbon. TPC and POC were subsequently measured on a mass spectrometer (ANCA-SL 20-20 Europa Scientific). Particulate inorganic carbon (PIC) was calculated as the difference between TPC and POC. Cell counts obtained with a Coulter Multisizer at the beginning and the end of incubations were used to calculate specific growth rates. PIC and POC production rates were calculated from cellular inorganic and organic carbon contents and specific growth rates.

Field

Ship-board productivity and calcification experiments were conducted at three stations in the subarctic North Pacific Ocean in June of 1998 (one experiment) and September of 1999 (four experiments). Station locations are given in the legend of Fig. 4. Surface seawater (10–20 m) was collected using a trace-metal-clean *in situ* pumping system and dispensed into acid-soaked polycarbonate bottles (3–4 replicate bottles per treatment). Samples were incubated on deck at about 30% surface irradiance levels in a flow-through incubator at *in situ* temperatures ($13 \pm 1^\circ\text{C}$). CO_2 concentrations in samples were manipulated by either bubbling with commercially prepared CO_2 /air mixtures (Station P26, 1998/1999) or by the addition of high-purity (trace-metal-clean) HCl/NaOH. Low CO_2 samples contained about $10 \mu\text{M CO}_2$ ($\sim 250 \text{ p.p.m.v.}$) with a pH of about 8.20 while high CO_2 treatments contained approximately $33 \mu\text{M CO}_2$ ($\sim 800 \text{ p.p.m.v.}$) with a pH of about 7.75. Total alkalinity ($\sim 2180 \mu\text{Eq l}^{-1}$) was unaffected by CO_2 bubbling while HCl and NaOH additions changed alkalinity by -6.4% and $+3.4\%$, respectively. CO_2

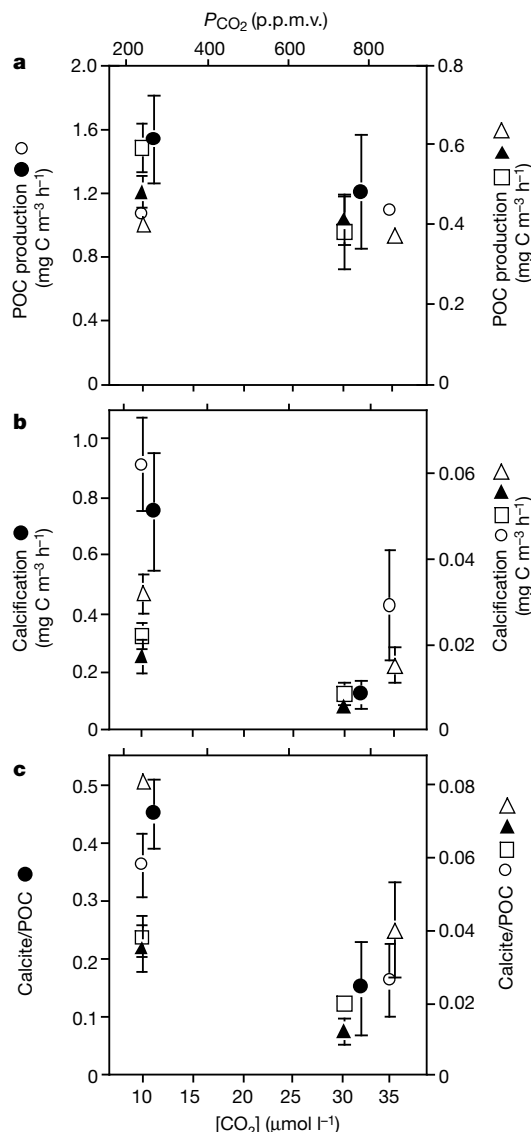


Figure 4 Effects of CO_2 manipulations on POC production, calcification and the ratio of calcification to POC production (calcite/POC) in subarctic North Pacific phytoplankton assemblages. **a**, POC production; **b**, calcification; and **c**, the calcite/POC ratio. Station P26 (50°N , 145°W) 1998, 6.8-day CO_2 pre-conditioning (filled circles). Station P26 1999, 2-day CO_2 pre-conditioning (squares). Station P26 1999, 9-day CO_2 pre-conditioning (filled triangles). Station P20 ($43^\circ 30' \text{N}$, $138^\circ 40' \text{W}$) 1999, 1.5-day CO_2 pre-conditioning (open triangles). Station Z9 (55°N , 145°W) 1999, 1.5-day pre-conditioning (open circles). In all five experiments, POC production did not differ significantly between CO_2 treatments (t -test, $p \geq 0.4$). The statistical significance of calcification rate differences (t -test) is as follows: P26, all experiments and measurements ($p < 0.05$); P20, calcification ($p = 0.056$); calcite/POC ($p = 0.074$); Z9, calcification ($p = 0.135$); calcite/POC ($p = 0.11$). Error bars represent standard errors of means.

concentrations in samples were calculated from measurements of pH and alkalinity using the algorithm developed by ref. 28. After a pre-conditioning period ranging from 1.5–9 days, samples were incubated with 40 $\mu\text{Ci } ^{14}\text{C}$ (50 mCi mmol⁻¹) for 6–9 hours, harvested onto 0.4- μm polycarbonate filters, and immediately frozen in scintillation vials at -20°C . In the laboratory, filter samples were acidified with either 10% HCl or H₃PO₄ to measure acid-stable (organic) carbon. The liberated acid-labile (particulate inorganic) carbon was trapped in NaOH solution contained either in small vials suspended in the primary vials or soaked into small GF/D glass fibre filters stuck onto the caps of the primary vials. Samples were measured on a liquid scintillation counter and corrected for ^{14}C uptake in dark control bottles as well as filtered seawater blanks.

Received 2 August 1999; accepted 18 July 2000.

1. Milliman, J. D. Production and accumulation of calcium carbonate in the ocean: budget of a nonsteady state. *Glob. Biogeochem. Cycles* **7**, 927–957 (1993).
2. Holligan, P. M. & Robertson, J. E. Significance of ocean carbonate budgets for the global carbon cycle. *Glob. Change Biol.* **2**, 85–95 (1996).
3. Houghton, J. T. et al. (eds) *Climate Change 1994: Radiative Forcing of Climate Change and an Evaluation of the IS92 Emission Scenario* (Cambridge Univ. Press, Cambridge, 1995).
4. Wolf-Gladrow, D. A., Riebesell, U., Burkhardt, S. & Bijma, J. Direct effects of CO₂ concentration on growth and isotopic composition of marine plankton. *Tellus* **51B**, 461–476 (1999).
5. Gattuso, J.-P., Frankignoulle, M., Bourge, I., Romaine, S. & Buddemeier, R. W. Effect of calcium carbonate saturation of seawater on coral calcification. *Glob. Planet. Change* **18**, 37–46 (1998).
6. Langdon, C., Takahashi, T., Sweeney, C., Chipman, D., Goddard, J., Marubini, F., Aceves, H., Barnett, H. & Atkinson, M. Effect of calcium carbonate saturation state on the calcification rate of an experimental coral reef. *Global Biogeochem. Cycles* **14**, 639–654 (2000).
7. Westbroek, P., Young, J. R. & Linschooten, K. Coccolith production (biomineralisation) in the marine alga *Emiliania huxleyi*. *J. Protozool.* **36**, 368–373 (1989).
8. Westbroek, P. et al. A model system approach to biological climate forcing. The example of *Emiliania huxleyi*. *Glob. Planet. Change* **8**, 27–46 (1993).
9. Winter, A., Jordan, R. W. & Roth, P. H. in *Coccolithophores* (eds Winter, A. & Siesser, W. G.) 161–179 (Cambridge Univ. Press, Cambridge, 1994).
10. Westbroek, P. et al. Strategies for the study of climate forcing by calcification. *Bull. Inst. Oceanogr. Monaco* (Spec. Issue) **13**, 37–60 (1994).
11. Holligan, P. M. et al. A biogeochemical study of the coccolithophore, *Emiliania huxleyi*, in the North Atlantic. *Glob. Biogeochem. Cycles* **7**, 879–900 (1993).
12. Nielsen, M. V. Growth, dark respiration and photosynthetic parameters of the coccolithophorid *Emiliania huxleyi* (Prymnesiophyceae) acclimated to different day length-irradiance combinations. *J. Phycol.* **33**, 818–822 (1997).
13. Young, J. R. Variation in *Emiliania huxleyi* coccolith morphology in samples from the Norwegian EHUX Experiment, 1992. *Sarsia* **79**, 417–425 (1994).
14. Booth, B. C., Lewin, J. & Postel, J. R. Temporal variation in the structure of autotrophic and heterotrophic communities in the subarctic Pacific. *Prog. Oceanogr.* **32**, 57–99 (1993).
15. Purdie, D. A. & Finch, M. S. Impact of a coccolithophorid bloom on dissolved carbon dioxide in sea water enclosures in a Norwegian fjord. *Sarsia* **79**, 379–387 (1994).
16. Frankignoulle, M., Canon, C. & Gattuso, J.-P. Marine calcification as a source of carbon dioxide: Positive feedback of increasing atmospheric CO₂. *Limnol. Oceanogr.* **39**, 458–462 (1994).
17. Morse, J. W. & Mackenzie, F. T. in *Developments in Sedimentology 48: Geochemistry of Sedimentary Carbonates* (Elsevier, Amsterdam, 1990).
18. Young, J. R. in *Coccolithophores* (eds Winter, A. & Siesser, W. G.) 63–82 (Cambridge Univ. Press, Cambridge, 1994).
19. McConnaughey, T. A. Calcification, photosynthesis, and global cycles. *Bull. Inst. Océanogr. Monaco* (Spec. Issue) **13**, 137–161 (1994).
20. McConnaughey, T. A. & Whelan, J. F. Calcification generates protons for nutrient and bicarbonate uptake. *Earth Sci. Rev.* **42**, 95–117 (1997).
21. Harris, R. P. Zooplankton grazing on the coccolithophorid *Emiliania huxleyi* and its role in inorganic carbon flux. *Mar. Biol.* **119**, 431–439 (1994).
22. Fritz, J. J. & Balch, W. M. A light-limited continuous culture study of *Emiliania huxleyi*: determination of coccolith detachment and its relevance to cell sinking. *J. Exp. Mar. Biol. Ecol.* **207**, 127–147 (1996).
23. Khesghi, H. S., Flannery, B. P. & Hoffer, M. I. Marine biota effects on the compositional structure of the world oceans. *J. Geophys. Res.* **96**, 4957–4969 (1991).
24. Guillard, R. R. L. & Rytner, J. H. Studies of marine planktonic diatoms. I. *Cyclotella nana* (Hustedt) and *Detonula confervacea* (Cleve). *Can. J. Microbiol.* **8**, 229–239 (1962).
25. Johnson, K. M., Wills, K. D., Butler, D. B., Johnson, W. K. & Wong, C. S. Coulometric total carbon dioxide analysis for marine studies: maximizing the performance of an automated gas extraction system and coulometric detector. *Mar. Chem.* **44**, 167–187 (1993).
26. Bradshaw, A. L., Brewer, P. G., Shafer, D. K. & Williams, R. T. Measurements of total carbon dioxide and alkalinity by potentiometric titration in the GEOSECS program. *Earth Planet. Sci. Lett.* **55**, 99–115 (1981).
27. Goyet, C. & Poisson, A. New determination of carbonic acid dissociation constants in seawater as a function of temperature and salinity. *Deep-Sea Res.* **36**, 1635–1654 (1989).
28. Lewis, E. & Wallace, D. W. R. Program Developed for CO₂ System Calculations. ORNL/CDIAC-105. (Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, Tennessee, 1998).

Acknowledgements

We thank A. Dauelsberg, B. Höhnisch, A. Terbrüggen, and K.-U. Richter for laboratory assistance, F. Hinz for REM analyses, D. Crawford, M. Lipsen, F. Whitney and C. Mayfield for invaluable help at sea and C. S. Wong for providing space on RV J.P. Tully. The *E. huxleyi* strain PML B92/11A was generously supplied by J. Green, Plymouth Marine Laboratory, and the *G. oceanica* strain PC 7/1 by the CODENET Algae collection in Caen. This work was supported by the Netherlands-Bremen Cooperation in Oceanography (NEBROC).

Correspondence and requests for materials should be addressed to U. R. (e-mail: uriebesell@awi-bremerhaven.de).

Increased dissolved oxygen in Pacific intermediate waters due to lower rates of carbon oxidation in sediments

Lowell D. Stott, William Berelson, Robert Douglas & Donn Gorsline

Department of Earth Sciences, University of Southern California, Los Angeles, California 90089, USA

Concentrations of dissolved oxygen in the ocean seem to correlate well with climate instabilities over the past 100,000 years. For example, the concentration of dissolved oxygen in Pacific intermediate waters was considerably higher during Pleistocene glacial periods than it is today^{1–4}. This has been inferred from the presence of bioturbated sediments, implying that oxygen levels were sufficient for burrowing organisms to live. Today, basins in the northeastern Pacific Ocean are floored by laminated sediments implying lower oxygen levels, which may be explained by reduced ventilation^{2–4}. Here we report a recent return to bioturbated sediments in the northeastern Pacific Ocean since the late 1970s. From the carbon isotope composition of benthic foraminifers living in the sediment, we infer a twofold decrease in the carbon oxidation rate occurring within sediments, equivalent to an increase in dissolved oxygen concentration of 15–20 micromoles per litre. These changes, at the edges of the Santa Barbara, Santa Monica and Alfonso basins, are coincident with a change in North Pacific climate which has reduced upwelling by 20–30% and increased sea surface temperatures by 1.5–3 °C. This suggests that climate effects on surface productivity, reducing the supply organic matter to sediments, may have had a greater effect on benthic oxygen levels than changes in ocean circulation patterns.

Laminated sediments have been accumulating in the Santa Barbara and Santa Monica basins, and basins within the Gulf of California, for the past few centuries. However, the spatial extent of laminated sediments within these basins has been changing⁵. In Santa Monica basin, a study of Holocene sediment patterns revealed that between AD 1600 and the 1970s, laminated sediments spread systematically outwards and upwards into shallower waters from the centre of the basin, so as to encompass the entire basin floor by the 1970s (ref. 5). During this same period, laminated sediments were also forming in Santa Barbara basin and in the basins along the margins of the Gulf of California^{6–11}. The temporal record of expanding laminated sediments in Santa Monica basin suggests that there are factors operating on centennial timescales that affect bottom-water oxygen levels, which probably control the distribution of bioturbating organisms. It is also now clear that there are shorter-timescale environmental changes that also affect the North Pacific marine environment, and that these are superimposed on the longer-scale patterns of variability. In particular, the mean climate state of the North Pacific, as measured by a number of different variables, has varied between a warm and a cold phase with a quasi-regular decadal oscillation^{12–19}. This decadal pattern has been referred to as the Pacific Decadal Oscillation¹⁹ (PDO). In the ocean, the PDO is characterized by a shift in the bifurcation point of the Sub-arctic Current as it approaches North America. This point of bifurcation strongly influences oceanographic, as well as weather patterns along the coast of the western USA^{15–18}. The last clear PDO shift was coincident with a 1976–77 El Niño event. Since that time, the average sea surface temperatures within the southern California Current during the months of upwelling (spring/summer) have increased by 1.5–3 °C (Fig. 1). Upwelling along the coast at the latitude of Santa Monica basin during the spring has also decreased (Fig. 1). The result of these changes has been a systematic decline in marine fisheries in the southern California Current,

which can be tied to the reduced upwelling and lower primary productivity¹².

An X-radiograph survey of box cores from Santa Barbara, Santa Monica and basins in the Gulf of California reveals that regional

pattern of laminated sedimentation was interrupted in the 1970s (Fig. 2). At that time bioturbated sediments began to replace laminated sediments along the edge of the laminated zone in each basin. The timing of this transition—determined from ²¹⁰Pb

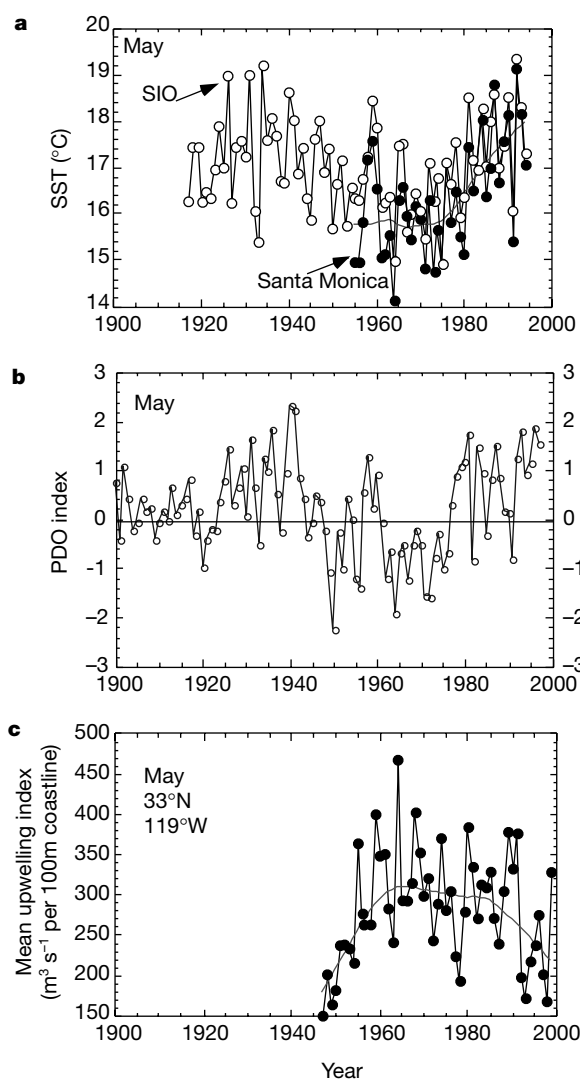


Figure 1 Climatic records from the northeastern Pacific. **a**, Sea surface temperature (SST) records collected at Scripps Pier and at Santa Monica. The smooth curve through the Santa Monica data is a locally weighted least-squared error method using a 50% population weighting. Both temperature records document the upward trend in temperatures that began in the late 1970s. **b**, May values of the PDO (Pacific decadal oscillation) index, which is a multivariate representation of environmental changes with

respect to the long-term average. The SST record is the first principal component in the PDO. We note the shift in the PDO from negative values to positive values in the late 1970s. **c**, Upwelling index for May at Santa Monica compiled by the Pacific Fisheries Environmental Laboratory for the years 1946–97. We note the increased upwelling trend before the late 1970s, when the trend reversed. The smoothed curve is as in **a**.

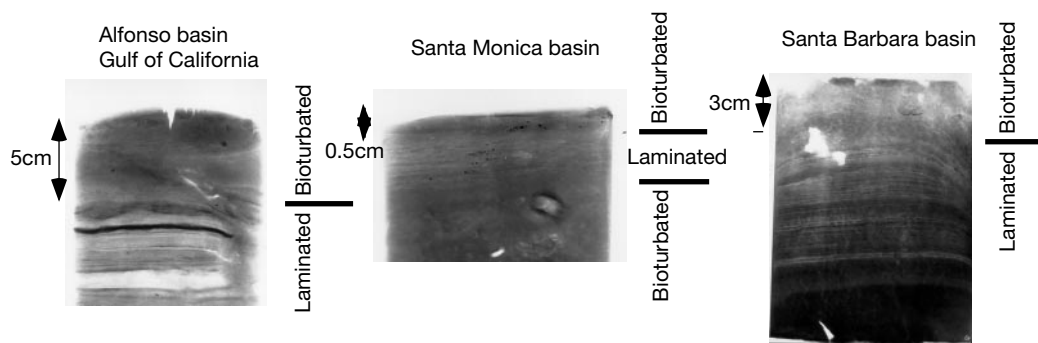


Figure 2 X-radiographs of sections through cores from the Santa Monica, Santa Barbara, and Alfonso basins. This figure shows the return of bioturbation in the topmost sediment (middle 1970s). Age estimates for the transition are extrapolated from a zero-age top,

using accumulation-rate estimates from ²¹⁰Pb-dated cores. The Santa Monica core was a companion core to the one used to develop the isotopic and mass-accumulation-rate data used in this study. It was collected at the edge of the basin floor.

dating—appears to be similar in each basin. However, the precise timing of the onset of bioturbation in the basins cannot be determined because bioturbation tends to smear the horizon between laminated and bioturbated sediments. The biological surveys that we have conducted in the Alfonso basin and the Borderland basins suggest that the type of bioturbation we observe in the topmost sediment horizons is produced by very small

organisms, such as polychaetes and perhaps some benthic foraminifera, that can tolerate very low oxygen conditions²⁰. These types of microbioturbators are largely restricted to the uppermost sediment horizons, near the sediment–water interface, where they produce horizontal disturbances; they do not extend deep into the sediment. This would explain why bioturbation extends to only 5 mm or so in the core from Santa Monica basin. The higher sediment accumulation rates characteristic of the Santa Barbara and Alfonso basins have produced a thicker package of bioturbated sediments in cores from these basins.

The return of bioturbated sedimentation in the 1970s means that there was an increased supply of oxygen to the basin deep waters, or that the rain rate (and oxidation rate) of organic carbon, which consumes oxygen, has been reduced. We have reviewed hydrographic records collected by the CALCOFI project for the time period 1960–90, and find no evidence of change in the amount of oxygen entering Santa Barbara basin at sill depth.

To assess whether changes of carbon oxidation within the basins could explain the apparent change in dissolved oxygen and bioturbation since the 1970s, we utilize two proxies for palaeoproductivity and carbon rain: the accumulation rates of organic carbon and carbonate carbon in the basins. In ²¹⁰Pb-dated multicores and box cores from the edge of the laminated zone in Santa Monica basin, the accumulation rate of organic carbon increased by approximately 20% between 1930 and 1970 (Fig. 3): the accumulation rate of carbonate carbon increased similarly. After the 1970s, the accumulation rates of both organic carbon and carbonate carbon decreased in concert (Fig. 3). Maximum organic-carbon values occur in the laminated intervals. A similar pattern of carbon and carbonate accumulation is observed in Santa Rosalia basin, although bioturbation extends deeper into cores from this basin. While some of the increased carbon accumulation observed between the 1950s and the 1970s could be explained by insufficient time for diagenetic ‘burn-down’, the recent decrease in carbon accumulation can be interpreted to reflect a drop in organic carbon and carbonate rain. Similarly, an earlier study of sediments from the Santa Barbara basin²¹ found that a decrease in the diatom flux started in the middle 1970s, reflecting a decline in coastal biological productivity that coincided with a warming of the California Current.

Several different processes could produce synchronous changes in the accumulation rates of organic carbon and carbonate carbon; these include variations in the rain of biogenic carbon, sediment accumulation and diagenesis^{22,23}. Our interpretation of ²¹⁰Pb profiles from cores in each basin suggests that the total mass accumulation rates have not changed since the 1970s (ref. 5, and L.D.S., W.B., R.D. and D.G., unpublished results). The recent decline of carbonate and organic carbon reflected in the Santa Monica basin since the late 1970s is a likely indicator of decreased carbon rain. The question is, can decreased carbon input and lower rates of carbon oxidation account for the apparent change in dissolved oxygen within basin deep waters that has promoted the re-initiation of bioturbation?

In the absence of larger macrofauna, the flux of CO₂ from the sediments into basin deep water is a diffusion-dominated process²², which varies as a function of the oxidation rate of organic carbon. Consequently, the gradients in total dissolved CO₂ (ΣCO₂) and porewater δ¹³C are proportional to the benthic ΣCO₂ flux. The porewater isotopic gradient is recorded in the carbon-isotope contrasts among benthic foraminifera that inhabit different horizons within the pore waters. As long as the species maintain a constant depth preference, the isotopic contrast will be sensitive to changes in the oxidation rate of organic carbon²⁴.

Two benthic foraminiferal species found in Santa Monica basin exhibit δ¹³C characteristics that suggest they are good recorders of the δ¹³C values of dissolved inorganic carbon of their respective habitats. *Bolivina argentea* inhabits the upper 1 mm of the sediment column, and *Buliminella tenuata* is most abundant at depths of 4 to

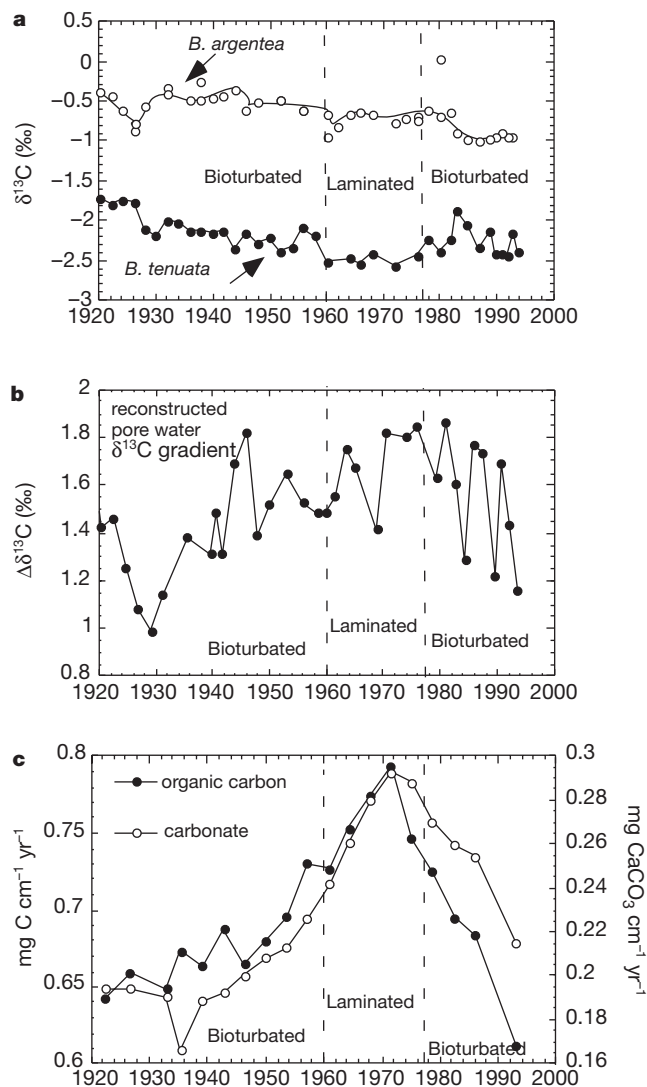


Figure 3 Reconstructed histories of porewater carbon-isotope gradient, and carbon accumulation rate, for the Santa Monica basin. **a**, Carbon-isotope composition of *B. argentea* (near-surface dweller) and *B. tenuata* (subsurface dweller) used to reconstruct porewater δ¹³C (ΣCO₂). A smooth curve is drawn through the data. In multicores and box cores from the Santa Monica and Santa Rosalia basins, the maximum abundance of living *B. argentea* occurs at 1 mm depth in the sediment; the abundance then drops to zero at 5 mm. The maximum abundance of *B. tenuata* is between 4 and 6 mm in both basins. Its live abundance drops off to zero by 9 mm. We infer that the horizon of maximum abundance corresponds to the location of preferred life habitat in the pore waters. This is supported by comparing the δ¹³C values of living specimens with the δ¹³C values of ΣCO₂ extracted from pore water at this site. The isotopic composition of stained specimens matches that of the pore water at the horizon of maximum abundance. **b**, Porewater carbon-isotope gradient reconstructed from the difference in δ¹³C of *B. argentea* and *B. tenuata* in samples collected from a multicore from 860 m water depth along the edge of the laminated zone in Santa Monica basin. We note the increased gradient before the 1970s, and then the reversal in the gradient as the sediments became bioturbated. **c**, Accumulation rates of organic carbon and carbonate carbon for the same Santa Monica core. We note how the values peaked in the 1970s, and then decreased towards the present.

6 mm. We observe the same depth distribution in the Gulf of California basins. The $\delta^{13}\text{C}$ values displayed by both species closely approximate that of ΣCO_2 extracted from pore fluids at the depth of their maximum abundance (and inferred calcification horizon; L.D.S. and R.D., unpublished results).

Between 1930 and 1970 the porewater isotopic gradient reconstructed from fossil foraminifera increased by nearly 1.0‰, following the trend in accumulation rates of organic and carbonate carbon (Fig. 3). The increased $\delta^{13}\text{C}$ gradient persisted until the late 1970s, when the gradient began to decrease and bioturbation began to replace laminated sediments. Taken together, the benthic foraminiferal $\delta^{13}\text{C}$ record and carbon accumulation rate data indicate that carbon oxidation changes were closely coupled to the changes in bottom-water oxygen that promoted changes in sedimentation.

We developed a simple reaction–diffusion model—to estimate how much change in carbon oxidation would be required to explain the observed changes in the porewater $\delta^{13}\text{C}$ gradient—by making three basic assumptions. We assume (1) a constant carbon isotopic value (equal to –20‰) for the organic matter that is oxidized (small changes in $\delta^{13}\text{C}$ do not affect the calculation), (2) that the mean residence time of deep water has not changed, and (3) that the relative depth offsets between the shallow- and deeper-dwelling infaunal benthic foraminifera remained constant. With these assumptions, the observed change in the benthic foraminifera $\delta^{13}\text{C}$ gradient would require a change of a factor of ~2 in the benthic carbon oxidation rate. An equivalent change in total oxygen consumption could alter the bottom-water oxygen concentrations by as much as 15–20 $\mu\text{mol l}^{-1}$. This would easily account for the observed shift from suboxic to oxic sediments.

The index of upwelling and the records of sea surface temperature for the southern California margin both indicate that primary productivity has diminished as a direct consequence of the climate shift that took place in the late 1970s. The close coupling between upwelling, sea surface temperature and carbon oxidation change that we document here highlights the importance of productivity in controlling bottom-water oxygen within the silled basins of the northeastern Pacific. These findings could have important implications for previous studies of oxygen variability in Pleistocene intermediate water—such variability has also been linked to climate change. On the basis of climate reconstructions by COHMAP²⁵, the northeastern Pacific margin experienced reduced northerly winds and reduced upwelling during the last glacial period: these changes in winds and upwelling might have influenced primary productivity and carbon rain rates. In light of the work that we report here, it is reasonable to consider whether the role of changing productivity may have been underestimated in previous studies of North Pacific intermediate water ventilation history. □

Received 11 May 1999; accepted 26 July 2000.

1. Keigwin, L. D. & Jones, G. A. Deglacial climatic oscillations in the Gulf of California. *Paleoceanography* **5**, 1009–1023 (1990).
2. Kennett, J. P. & Ingram, B. L. A 20,000-year record of ocean circulation and climate change from the Santa Barbara Basin. *Nature* **377**, 510–514 (1995).
3. Behl, R. J. & Kennett, J. P. Evidence for brief interstadial events in the Santa Barbara Basin, NE Pacific during the past 60 kyr. *Nature* **379**, 243–246 (1996).
4. Stott, L. D., Neumann, M. & Hammond, D. Intermediate water ventilation on the northeastern Pacific margin during the late Pleistocene inferred from benthic foraminiferal $\delta^{13}\text{C}$. *Paleoceanography* **15**, 161–169 (2000).
5. Christensen, C. J., Hammond, D. E. & Lund, S. P. Non-annual laminations and expansion of anoxic basin-floor conditions in Santa Monica Basin, California Borderland, over the past four centuries. *Mar. Geol.* **116**, 399–418 (1994).
6. Nava-Sanchez, E., Gorsline, D., Douglas, R., Rikansrud, K. & DeDiego, T. Paleoclimatic/paleoceanographic studies of Gulf of California margin basin, Baja California Sur, Mexico. *Eos* **F79** (abstr.) (1997).
7. Alvarez-Borrego, S. & Lara-Lara, J. in *The Gulf and Peninsular Province of the Californias* (ed. Dauphin, J. & Simoneit, B.) 555–568 (Memoir 47, American Association of Petroleum Geologists, Tulsa, 1991).
8. Douglas, R., DeDiego, T., Gorsline, D. & Nava, E. Modern “Black Shales” in the Gulf of California. *Abstract Geol. Soc. Am.* **30**, A-55 (1998).
9. DeDiego, T. & Douglas, R. Oxygen-related biofacies in black shales from the Gulf of California, Mexico. *W. V. Sliter Memorial Symposium* (eds Bralower, T. & Huber, B.) *J. For. Res.* (in the press).
10. Soutar, A. & Crill, P. A. Sedimentation and climatic patterns in the Santa Barbara Basin during the 19th and 20th centuries. *Geol. Soc. Am. Bull.* **88**, 1161–1172 (1977).

11. Biondi, F., Lange, C. B., Hughes, M. K. & Berger, W. H. Inter-decadal signals during the last millennium (1117–1992) in the varve record of Santa Barbara basin, California. *Geophys. Res. Lett.* **24**, 193–196 (1997).
12. Roemmich, D. & McGowan, J. Climatic warming and the decline of zooplankton in the California Current. *Science* **267**, 1324–1326 (1995).
13. Venrick, E. L., McGowan, J. A., Cayan, D. R. & Hayward, T. L. Climate and chlorophyll *a*: Long-term trends in the central North Pacific Ocean. *Science* **238**, 70–72 (1987).
14. Namias, J. Multiple causes of the North American abnormal winter of 1976–1977. *Mon. Weath. Rev.* **106**, 279–295 (1978).
15. Trenberth, K. E. Recent observed interdecadal climate changes in the northern hemisphere. *Bull. Am. Meteorol. Soc.* **71**, 988–993 (1990).
16. Ebbesmeyer, C. C., Coomes, C. A., Cannon, G. T. A. & Bretschneider, D. E. in *Climate Variability on the Eastern Pacific and Western North America* (ed. Peterson, D. H.) 399–417 (Geophysical Monograph 55, American Geophysical Union, Washington DC, 1989).
17. Graham, N. E. Decadal-scale climate variability in the 1970s and 1980s: observations and model results. *Clim. Dyn.* **10**, 135–159 (1994).
18. Trenberth, K. E. & Hurrell, J. W. Decadal atmosphere-variations in the Pacific. *Clim. Dyn.* **9**, 303–319 (1994).
19. Mantua, N. J., Hare, S. R., Zhang, Y., Wallace, J. M. & Francis, R. C. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bull. Am. Meteorol. Soc.* **78**, 1069–1079 (1997).
20. Levin, L. A., Gage, J. D., Martin, C. & Lamont, P. A. Macrobenthic community structure within and beneath the oxygen minimum zone, NW Arabian Sea. *Deep-Sea Res.* **47**, 189–226 (2000).
21. Lange, C. B., Burke, S. K. & Berger, W. H. Biological production off southern California is linked to climate change. *Clim. Change* **16**, 319–329 (1990).
22. Jahnke, R. A. Early diagenesis and recycling of biogenic debris at the sea floor, Santa Monica Basin, California. *J. Mar. Res.* **48**, 413–438 (1990).
23. Berelson, W. M. *et al.* Biogenic matter diagenesis on the sea floor: A comparison between two continental margin transects. *J. Mar. Res.* **54**, 731–762 (1996).
24. McCorkle, D. & Emerson, S. R. The relationship between pore water carbon isotopic composition and bottom water oxygen concentration. *Geochim. Cosmochim. Acta* **52**, 1169–1178 (1988).
25. COHMAP. Climatic changes of the late 18,000 years: Observations and model simulations. *Science* **241**, 1043–1052 (1988).

Acknowledgements

This work was supported by the NSF.

Correspondence and requests for materials should be addressed to L.D.S. (e-mail: stott@usc.edu).

Variance in ecological consumer–resource interactions

Lisandro Benedetti-Cecchi

Dipartimento di Scienze dell’Uomo e dell’Ambiente, Via A. Volta 6, I-56126 Pisa, Italy

Food-web models use the effect size of trophic interactions to predict consumer–resource dynamics^{1–3}. These models anticipate that strong effects of consumers increase spatial and temporal variability in abundance of species, whereas weak effects dampen fluctuations^{4–6}. Empirical evidence indicates that opposite patterns may occur in natural assemblages⁷. Here I show that spatial variance in the distribution of resource populations is sensitive to changes in the variance of the trophic interaction, in addition to the mean effect of consumers, relative to other causes of spatial variability. Simulations indicate that both strong and weak direct effects of consumers can promote spatial variability in abundance of resources, but only trophic interactions with a large mean effect size can reduce variation. Predictions of the model agree with the results of repeated field experiments and are consistent with data from published consumer–resource interactions, proving to be robust across widely varying environmental conditions and species’ life histories. Thus, food-web models that embody variance in trophic interactions may have increased capacity to explain the wide range of effects of consumers documented in empirical studies.

Consumers may affect both the mean abundance and the spatial and/or temporal variance of their resources^{7,8}. Effects on the

variance reflect both the intrinsic variability of the trophic interaction and concomitant changes due to the mathematical relationship between the mean and the variance⁹. This relationship is dome-shaped for populations whose abundance is expressed as percentage cover. The largest variances correspond to mean values around 50%, whereas variances decrease for mean values approaching 0 and 100%. This should also hold for populations whose abundance is expressed as number of individuals per unit of surface, with variances decreasing as mean abundance approaches 0 or the density at which space is saturated.

I derived a conceptual model from these basic principles (Fig. 1), to predict changes in spatial variance of a resource as a function of the variance and mean effect of the trophic interaction, the residual variability of the resource (that is, the component of variation that is not due to the consumer) and the overall abundance attainable by the resource in the absence of the consumer (sparse versus abundant populations). The rationale is that a trophic interaction is expected to increase spatial variation in a resource if the variance of the trophic interaction is larger than residual variation, and if the amount of variability generated by the consumer is larger than the loss of variance that might follow the reduction in mean abundance of the resource. This, in turn, is a function of the density

of the resource population and the mean effect size of the trophic interaction.

The model predicts no significant effect on variation in abundance of the resource for weak trophic interactions with small variance compared to the residual variability of the resource, whether this is a sparse or an abundant population (Fig. 1a). A large mean effect size of a consumer, combined with a small variance of the trophic interaction, causes a strong reduction in spatial variance of sparse populations (Fig. 1b). This pattern is largely driven by the dampening effect on the variance that results from the strong reduction in mean abundance of the resource. This effect is not expected for a dominant population because its final abundance will be larger than that attained by a sparse population for any effect size of the consumer (expressed as proportional change), so that large fluctuations in abundance are still possible. If the effect of the consumer is small on average, but the variance of the trophic interaction is larger than residual variability (Fig. 1c), an increase in spatial variability of the resource is expected regardless of its abundance. Finally, consumers should increase the variability of their resources when both the mean effect and the variance of the trophic interaction are large (Fig. 1d). This should be more pronounced when the resource is a dominant population because its larger size will allow for larger fluctuations in abundance for any effect size of the consumer. Other possible situations (for example, variable effects of consumers with large residual variability in the resource) were not investigated because they were not directly predictable from these basic principles.

I used Monte Carlo simulations to examine the predictions of the conceptual model in a simple system, and to expand the model in order to include cases that were not explicitly addressed in its first formulation. Simulations explored the sensitivity of small-scale (metres) spatial heterogeneity of a resource (ephemeral algae) to changes in both the mean effect of consumers (limpets) and the variance of the trophic interaction, under different regimes of environmental variation and for resource populations of different size. The Monte Carlo procedure simulates a factorial experiment where limpets are excluded from patches of substratum in several sites at different times, based on the following linear model:

$$y_{ijk} = \mu + S_i + T_j + L_k + ST_{ij} + SL_{ik} + TL_{jk} + STL_{ijk} + \epsilon_{r(ijk)} \quad (1)$$

where y_{ijk} is the n th observation within any given combination of the other factors, μ is the general mean of the population, S_i is the effect of site i ($i = 4$), T_j is the effect of time j ($j = 4$), L_k is the effect of the k th level of limpets ($k = 2$, presence versus absence), $\epsilon_{r(ijk)}$ is the residual term associated with any given observation ($r = 3$ replicates), and other terms indicate the interactions among specified effects. I used equation (1) to simulate the effects of grazing in midshore habitats of rocky coasts under realistic environmental conditions. Grazing is a random variable and has a variance component; this is different from what is done in field experiments where grazing is usually examined as a fixed effect (see Methods). I repeated simulated experiments for different combinations of residual variability, mean and variance of the effect of consumers, and for abundant and sparse populations. A simulation consisted of 1,000 experiments for each of the conditions above. A single experiment provided an estimate of the effect of limpets on small-scale spatial variance of ephemeral algae. For simplicity, estimates were averaged across sites and times. I calculated the effect size for spatial variance of the resource as $\ln(\sigma_{+C}^2/\sigma_{-C}^2)$, where σ_{+C}^2 is the variance among replicates in the presence of consumers (in a particular site at a particular time), and σ_{-C}^2 is the variance in the absence of consumers¹⁰.

Results of the Monte Carlo simulations supported the qualitative predictions of the conceptual model (Fig. 2). Variable trophic interactions enhanced spatial heterogeneity of abundant (Fig. 2a) and sparse (Fig. 2d) resource populations with low residual variability, regardless of whether the mean effect size of consumers was

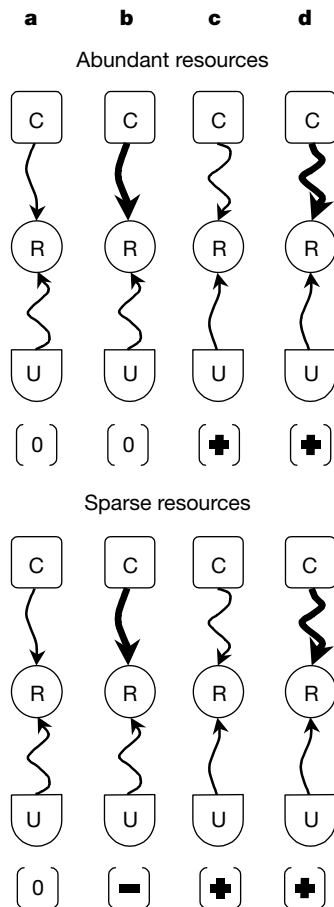


Figure 1 Conceptual model of the effect of consumers on variation in the abundance of a resource. C, Consumers; R, resource. **a–d**, different combinations of variance and mean effect of the trophic interaction and residual variability for sparse and abundant populations. The trophic interaction is represented by an arrow from the consumer to the resource, while residual variation is represented by an arrow from U (the processes responsible for residual variation) to R. Thickness of arrows is proportional to the mean effect size of the interaction, while the amplitude of the waves is proportional to the variance. The predicted effects are: no effect (0), decrease (–), and increase (+) in spatial variance of the resource.

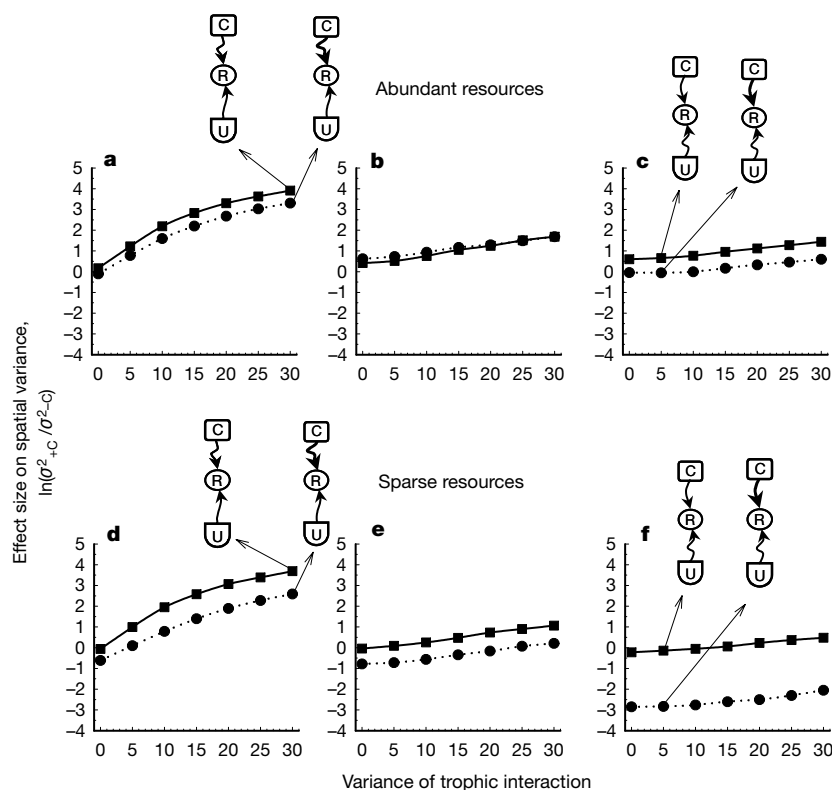


Figure 2 Results of Monte Carlo simulations illustrating the effects of the variance of trophic interactions. These were run with large (circles) and small (squares) mean effect size on small-scale spatial variance of the resource, under different combinations of mean

abundance and residual variability of the resource: low (**a, d**), intermediate (**b, e**) and large (**c, f**) residual variability. Note that a difference of 0.69 between large and small mean effect sizes corresponds to a 100% increase in spatial variance ($\ln 2 = 0.69$).

large or small. Weak interactions were generally more effective in enhancing spatial variation, although this was more evident for sparse (Fig. 2d, e) than abundant (Fig. 2a–c) resources. Consumers that had a large mean effect on a sparse resource with large residual variance strongly depressed spatial heterogeneity (Fig. 2f). A small, but variable effect size of consumers resulted in increased spatial variance of the resource at intermediate levels of residual variability, both for abundant (Fig. 2b) and sparse (Fig. 2e) populations. A similar pattern occurred for strong trophic interactions, but only when the resource was abundant; a reduction in spatial variation resulted for a sparse resource under steady trophic interactions (Fig. 2e).

To test whether the model above provided realistic predictions, I experimentally excluded/included limpets from patches of substratum at mid-shore heights of rocky coasts in the northwest Mediterranean. The experiment was replicated at 48 sites established at three locations (hundreds of kilometres apart), in two different seasons (summer and winter), on two dates within each season and on substrata of different slope (horizontal versus vertical). Empirical estimates of the effects of limpets on small-scale spatial variance of ephemeral algae were compared with those predicted by the model (see Methods). By repeating the experiment many times under different environmental settings, it was possible to compare empirical data with predictions of the model for several combinations of the simulated parameters. As an additional test I compared predictions of the model with data assembled from the literature, which included experimental analyses of the effects of grazing^{11–21} and predation^{22–24} in rocky intertidal assemblages. The predicted and observed effects of consumers were significantly correlated in both analyses (Fig. 3a, b). Thus, the model provides realistic predictions for consumer–resource interactions under a wide range of environmental conditions and species' life histories.

Recent empirical studies have shown that weak effects of consumers can amplify spatial variation in the distribution of resource populations, while strong interactions dampen fluctuations⁷. These patterns contrast with the predictions of several models of consumer–resource interactions^{1–3}. My results show that the response of a resource population to the presence of a consumer is not simply a function of the mean effect size of the trophic interaction, as implied by many food-web models. Both strong and weak trophic interactions can increase or decrease spatial variation in a resource, depending on the variance of the interaction, the abundance of the resource and its residual variability. Food-web and interaction-web models that embody the variance of the trophic interaction, weighed over the residual variability of the resource, may have increased capacity to explain the wide range of consumer–resource interactions documented in empirical studies. The significant relationship between predicted and observed effect sizes of consumers reported in the present study supports this conclusion.

Understanding variance in consumer–resource interactions and the underlying mechanisms (for example, how environmental factors affect spatial variation in density and in foraging efficiency of consumers^{25–27}), has important ramifications in a variety of ecological contexts, including the refinement of theories on the stability of both consumer and resource populations^{4,5,28}, the distinction between keystone predation and context-dependent effects of consumers^{22–25}, and the prediction of the effects of environmental change²⁶. For example, global climate change is expected to enhance the residual variability of many natural populations due to increased frequency in the occurrence of extreme disturbance events such as storms²⁹. In this situation, my results predict that weak trophic interactions will become ineffective in influencing spatial variation in abundance of both dominant (Fig. 2c) and sparse (Fig. 2f) resource populations. Given the importance of consumers in natural assemblages, the capability to anticipate

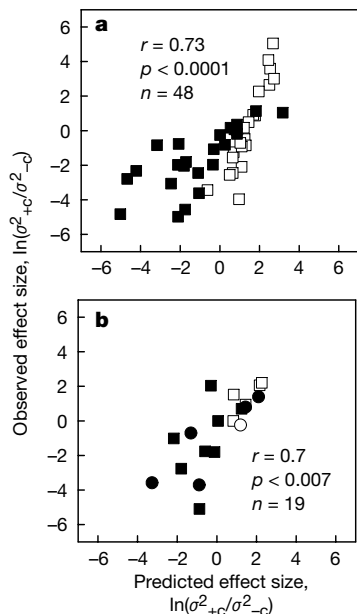


Figure 3 Correlation between observed and predicted effects of consumers on spatial variance of their resource. **a**, Results from experimental manipulations of limpets in the northwest Mediterranean. **b**, Results from an analysis of published data on the effects of grazers (squares) and predators (circles) in several rocky intertidal habitats; consumers affect resource populations with low (5–80, unfilled symbols) and high (100–750, filled symbols) residual variance. A study on algae–limpet interactions illustrates how observed and predicted effects were obtained¹¹. The variance among replicates in algal abundance was 392 in the presence of the consumer and 50 in its absence, and the mean cover was 53% in grazed plots and 92% in ungrazed plots. From these data, the observed effect size on the variance was calculated as $\ln(392/50) = 2.06$. The predicted effect size is obtained from equation (2) as a function of the variance in algal cover in ungrazed plots (50); the pooled variance among consumer-present and consumer-absent plots $((392 + 50)/2 = 221)$; and the effect size on mean percentage cover of the resource $(\ln(53/92) = -0.55)$. The predicted effect size for this example is 2.16.

which trophic interactions will be relevant under modified climate conditions is an important issue for environmental conservation and management. □

Methods

Monte Carlo simulations

Levels for the factors site, time and all their interactions were obtained by sampling normal distributions with mean zero and variance equal to the variance estimated from real data. These data came from published experimental studies on effects of grazing and from analyses of spatial and temporal patterns of abundance of ephemeral algae in the northwest Mediterranean^{27,30}. Levels for the effect of consumers were chosen by sampling normal distributions with means that simulate reductions in mean abundance of the resource to 0.9 (small effect size) and 0.2 (large effect size) times its original value in the presence of consumers, and variances in the range 0–30. Residuals were generated using an exponential distribution to simulate low (10), mid (200) and large (600) variation in the data. Mean percentage cover values in the absence of consumers were 35% (sparse population) and 70% (abundant population). Each simulated experiment provided estimates of the variance among replicates in the presence and in the absence of consumers, the residual variation, the effect size on mean percentage cover and the effect size on small-scale spatial variance of the resource. The effect size on mean percentage cover is calculated as $\ln(+C/-C)$, where $+C$ is the mean percentage cover of ephemeral algae in the presence of consumers and $-C$ is the mean cover in the absence of consumers (see above for the definition of the effect size on spatial variance).

Multifactorial experiment

The experimental design included two replicate dates within each combination of season and location, and two replicate sites (10–20 m stretches of the coastline) for each combination of inclination of the substratum, date, season and location. Thus, a total of 16 sites was established along 4–5 km of shoreline at each location. Nine plots of 17×17 cm were established in each site in mid-shore areas occupied by limpets and barnacles and

with minor or no cover of ephemeral algae. These plots were randomly distributed in triplicates among the following treatments: (1) control (plots were marked with epoxy putty for relocation and left undisturbed thereafter); (2) inclusion (limpets were maintained in the plots at the same density as the controls); (3) exclusion (plots were maintained free of limpets). Limpets were excluded/included using fences made of a plastic mesh reinforced with a mesh of galvanized iron. Fences were $17 \times 17 \times 4$ cm in size (mesh size of 0.5×0.5 cm), and were anchored to the substratum with stainless-steel screws inserted into rawl-plugs in the rock. The experiment started in June 1996 and each site was visited every 2–3 months to maintain the experimental conditions. The percentage cover of algae was assessed in areas of 10×10 cm in the middle of each plot as intercepts on a grid of 64 points. I used data after one year from the start of the experiment for analyses. Because of evidence of artefacts caused by the fences at some sites, only data from inclusion and exclusion plots were compared here. The variance among replicates was calculated for each site and used to estimate the effect size of limpets on small-scale spatial variation of ephemeral algae.

Analysis of published data

Estimates of the effects of consumers (grazers and carnivores) on spatial variation of their resources were obtained from 14 studies^{11–24}, chosen to represent trophic interactions on rocky shores under a wide range of environmental conditions. Studies that included several independent experiments provided more than one estimate of effect size. When data were derived from time series only one date (chosen at random) was used, to ensure independence of the data.

Predicted versus observed measures of effect size

I used multiple regression on data from the simulation experiment to predict the effect of consumers on the spatial variance of the resource (dependent variable) as a function of the following predictor variables: *a*, variance in the abundance of the resource in the absence of consumers; *b*, the pooled variance among consumer-present (+) and consumer-absent (–) plots; and *c*, the effect size on the mean percentage cover of the resource. Variance of the trophic interaction was not included because trophic interactions are examined as fixed effects in real experiments. The resulting equation

$$y = 1.46 - 0.0108a + 0.0068b + 0.4764c \quad (2)$$

($R^2 = 0.86$) was used to predict the effect of consumers on the spatial variance of the resource from field experiments. Predicted values were compared with observed estimates calculated directly from variances among replicates in the presence and in the absence of consumers, as described above.

Received 3 March; accepted 5 July 2000.

1. Lawton, J. H. in *Ecological Concepts* (ed. Cherrett, J. M.) 43–78 (Blackwell Scientific, Oxford, 1989).
2. Paine, R. T. Food webs: linkage, interaction strength and community infrastructure. *J. Anim. Ecol.* **49**, 667–685 (1980).
3. Paine, R. T. Food-web analysis through field measurement of per capita interaction strength. *Nature* **355**, 73–75 (1992).
4. May, R. M. *Stability And Complexity In Model Ecosystems* (Princeton Univ. Press, Princeton, 1973).
5. Polis, G. A. & Strong, D. R. Food web complexity and community dynamics. *Am. Nat.* **147**, 813–846 (1996).
6. McCann, K., Hastings, A. & Huxel, G. R. 1998. Weak trophic interactions and the balance of nature. *Nature* **395**, 794–798 (1998).
7. Berlow, E. L. Strong effects of weak interactions in ecological communities. *Nature* **398**, 330–334 (1999).
8. Fairweather, P. G. Predation can increase variability in the abundance of prey on seashores. *Oikos* **53**, 87–92 (1988).
9. Taylor, R. L. Aggregation, variance and the mean. *Nature* **189**, 732–735 (1961).
10. Osenberg, C. W., Sernelle, O. & Cooper, S. D. Effect size in ecological experiments: the application of biological models in meta-analysis. *Am. Nat.* **150**, 798–812 (1997).
11. Sousa, W. P. Intertidal mosaics: patch size, propagule availability, and spatially variable patterns of succession. *Ecology* **65**, 1918–1935 (1984).
12. Dungan, M. L. in *Predation: Direct And Indirect Impacts On Aquatic Communities* (eds Karefoot, W. C. & Sih, A.) 188–220 (Univ. Press New England, Hanover, 1987).
13. Martin, T. H., Wright, R. A. & Crowder, L. B. Non-additive impact of blue crabs and spot on their prey assemblages. *Ecology* **70**, 1935–1942 (1989).
14. Dethier, M. & Duggins, D. Variation in strong interactions in the intertidal zone along a geographical gradient: a Washington-Alaska comparison. *Mar. Ecol. Prog. Ser.* **50**, 97–105 (1988).
15. Menge, B. A., Lubchenco, J., Gaines, S. D. & Ashkenas, L. R. A test of the Menge-Sutherland model of community organization in a tropical rocky intertidal food web. *Oecologia* **71**, 75–89 (1986).
16. Lubchenco, J. *et al.* Structure, persistence and role of consumers in a tropical rocky intertidal community (Taboguilla Island, Bay of Panama). *J. Exp. Mar. Ecol.* **77**, 23–73 (1984).
17. Berlow, E. L. & Navarrete, S. A. Spatial and temporal variation in rocky intertidal community organization: lessons from repeating field experiments. *J. Exp. Mar. Biol. Ecol.* **214**, 195–229 (1997).
18. Sousa, W. Experimental investigations of disturbance and ecological succession in a rocky intertidal algal community. *Ecol. Monogr.* **49**, 227–254 (1979).
19. Dungan, M. L. Three-way interactions: barnacles, limpets, and algae in a sonoran desert rocky intertidal zone. *Am. Nat.* **127**, 292–316 (1986).
20. Morrison, D. Comparing fish and urchin grazing in shallow and deeper coral reef algal communities. *Ecology* **69**, 1367–1382 (1988).
21. Anderson, M. J. & Underwood, A. J. Effects of gastropods grazers on recruitment and succession of an estuarine assemblage: a multivariate and univariate approach. *Oecologia* **109**, 442–453 (1997).

22. Navarrete, S. A. Variable predation: effects of whelks on a mid-intertidal successional community. *Ecol. Monogr.* **66**, 301–321 (1996).
23. Wootton, T. J. Size-dependent competition: effects on the dynamics vs. the end point of mussel bed succession. *Ecology* **74**, 195–206 (1993).
24. Lubchenco, J. & Menge, B. A. Community development and persistence in a low rocky intertidal zone. *Ecol. Monogr.* **48**, 67–94 (1978).
25. Menge, B. A. *et al.* The keystone species concept: variation in interaction strength in a rocky intertidal habitat. *Ecol. Monogr.* **64**, 249–286 (1994).
26. Sanford, E. Regulation of keystone predation by small changes in ocean temperature. *Science* **283**, 2095–2097 (1999).
27. Benedetti-Cecchi, L. Predicting direct and indirect interactions during succession in a midlittoral rocky shore assemblage. *Ecol. Monogr.* **70**, 45–72 (2000).
28. Underwood, A. J. in *Frontiers of Population Ecology* (eds Floyd, R. B., Sheppard, A. W. & De Barro, P. J.) 369–389 (CSIRO, Melbourne, 1996).
29. Smith, S. V. & Buddemeier, R. W. Global change and coral reef ecosystems. *Annu. Rev. Ecol. Syst.* **23**, 89–118 (1992).
30. Menconi, M., Benedetti-Cecchi, L. & Cinelli, F. Spatial and temporal variability in the distribution of algae on rocky shores in the northwest Mediterranean. *J. Exp. Mar. Biol. Ecol.* **233**, 1–23 (1999).

Acknowledgements

I thank M. J. Anderson and F. Micheli for helpful comments on the manuscript. F. Bulleri, I. Bertocci and M. Menconi provided invaluable assistance in the field. This research was supported by the European Community and by a grant from the University of Pisa.

Correspondence and requests for materials should be addressed to L.B.-C. (e-mail: bencecc@discat.unipi.it).

Body image as a visuomotor transformation device revealed in adaptation to reversed vision

Kaoru Sekiyama*, Satoru Miyauchi†, Toshihide Imaruoka‡, Hiroyuki Egusa§ & Takara Tashiro||

* Division of Cognitive Psychology, Future University-Hakodate, Kameda Nakano, Hakodate 041-8655, Japan

† Auditory & Visual Informatics Section, Kansai Advanced Research Center, Communications Research Laboratory, Iwaoka, Kobe 651-2492, Japan

‡ Graduate School of Medicine, Osaka University, Yamadaoka, Suita 565-0781, Japan

§ School of Health, International University of Health and Welfare, Ohtawara 324-8501, Japan

|| Department of Psychology, Osaka City University, Sugimoto, Osaka 558-8585, Japan

People adapt with remarkable flexibility to reversal of the visual field caused by prism spectacles^{1,2}. With sufficient time, this adaptation restores visually guided behaviour and perceptual harmony between the visible and tactile worlds^{1–3}. Although it has been suggested that seeing one's own body is crucial for adaptation^{1,2}, the underlying mechanisms are unclear. Here we show that a new representation of visuomotor mapping with respect to the hands emerges in a month during adaptation to reversed vision. The subjects become bi-perceptual^{3–5}, or able to use both new and old representations. In a visual task designed to assess the new hand representation, subjects identified visually presented hands as left or right by matching the picture to the representation of their own hands. Functional magnetic resonance imaging showed brain activity in the left posterior frontal cortex (Broca's area) that was unique to the new hand representations of both hands, together with activation in the intraparietal sulcus and prefrontal cortex. The emergence of the new hand representation coincided with the adaptation of perceived location of visible objects in space. These results suggest that the hand

representation operates as a visuomotor transformation device that provides an arm-centred frame of reference⁶ for space perception.

Two male and two female right-handed volunteer psychology students (aged 20–24) served as subjects. They wore left–right reversing spectacles continuously for 39 (two males) or 35 (two females) days. The visual field was about 70° horizontally with 20° binocular overlap, and 60° vertically.

Although visually guided behaviour was disrupted at the beginning of prism wearing, all the subjects could ride a bicycle after one month of adaptation (see Supplementary Information). We examined visually guided hand movements by asking the subjects to reach a visual target presented at various positions for 500 ms. Both error and latency returned to the pre-exposure level within 2 weeks (Fig. 1a), indicating relatively fast adaptation. In contrast, restoration of the perceived location of visible objects was much slower. A visual target was presented on either the left or right side of a display. During the first 2 weeks, the subjects could not correctly report on which side the target was displayed (Fig. 1b). Although they regained correct localization by day 24, reaction time (RT) indicated a slower adaptation for the left (non-preferred hand) side even on day 35. This time course of perceptual adaptation coincided with the emergence of new body image described below.

We examined the subjects' body image (activated state of body representation) by asking them to identify pictures of hands viewed from different angles as a left or right hand (Fig. 2i). Usually, subjects solve this task by mentally moving an image of their own hand from its 'home position' into a match with the visual stimulus^{7–10}. We expected that the subjects' responses would be incorrect at the beginning of prism wearing because the prism reverses the appearance of stimuli, and that correct responses would occur as a new hand representation was established in long-term memory (Fig. 3a, b, c).

Before wearing the prism, responses were almost 100% correct (Fig. 2a). Early in prism wearing, correct responses essentially disappeared (Fig. 2b, c). However, after 3 weeks of adaptation, correct responses reappeared, suggesting the emergence of a new

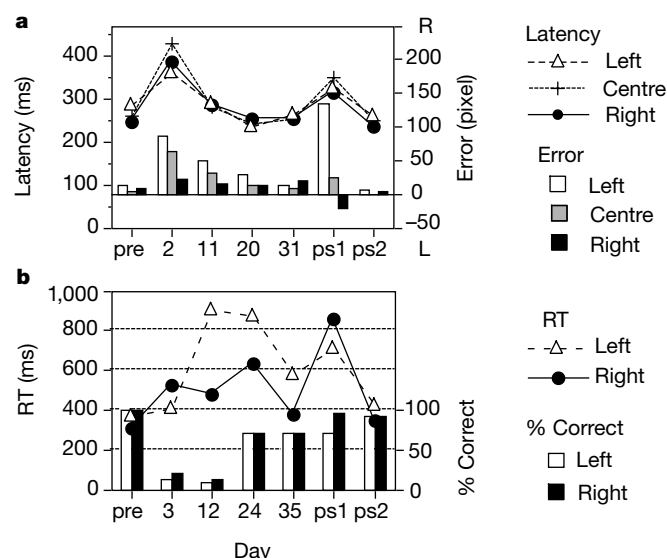


Figure 1 Performance in the reaching and visual localization task. **a**, Mean latency and reaching error for left, centre and right targets. Reaching error was defined as the horizontal distance between the target and reached position (pixels on a 20-inch display with 640 × 400 resolution). Positive errors indicate reaching to the right of targets, negative errors indicate reaching to the left. One subject was excluded from averaging owing to a recording failure. **b**, Mean reaction time (RT) and per cent correct localization for the left and right targets. Ps1: post1, within an hour of the removal of the reversing spectacles. Ps2: post2, the day after removal.

hand representation. On day 25, there were more correct responses for right-hand stimuli, especially in prototypical orientations ('home position'⁸) around 0° or 315° (Fig. 2d). This result may indicate that the new hand representation is generated earlier for the preferred hand in its most familiar orientation and that the new representation is most easily activated for matching when the stimulus orientation is close to it. At this point, the subjects reported that they could visualize their own hands in two ways: in the old relationship with somatic cues (Fig. 3b), and in the reversed relationship (Fig. 3c), suggesting that the old and new representations coexisted.

Despite the double representations, the subjects could selectively use the new pair if necessary. On day 25, another session of the hand identification task was given with the instruction to "use only your new hand image in all the trials". The subjects could follow this instruction, producing mostly correct responses (Fig. 2f). Although the RT became much longer, the increase was systematic, indicating that the new representation was acquired earlier for the preferred hand in its prototypical orientation. The steep slope of the RT function would represent awkwardness in mentally moving the novel hand image at this stage (see Fig. 2g for a later stage).

At the final stage of exposure, the RT function was different from that in the pre-test, showing increased RTs for stimuli around 0° (Fig. 2e). Additional data (Fig. 2g, h) indicate that the increase may be mainly due to conflict between the new and old hand images at these familiar orientations, perhaps because the instruction emphasized attention to both new and old images. The subjects reported that the task would have been much simpler at this stage (Fig. 2e) if they had been allowed to concentrate on either the new or old images, as they were during the functional magnetic resonance imaging (fMRI) trials (Fig. 2g, h).

To assess the neuronal differences between the old and new hand representations, we scanned each subject's whole brain using fMRI at the final stage of adaptation (days 34–38). We tested the subjects using the hand identification task and explicitly instructed them to use either the new or old hand image only. In each instruction condition, there were also two stimulus identity conditions, a left-hand condition (a left hand was presented in 81% of the trials) and a right-hand condition (a right hand was presented in 81% of the trials).

For each of the four conditions, we analysed the fMRI signals during the test periods against those during control periods. The

areas that were active for all of the conditions in common were the parietal association cortex (Brodmann's area (BA) 7 or superior parietal lobule), premotor cortex (BA6), visual association cortex (BA18, 19, 37) and cerebellum (Fig. 4a–d). These areas are responsible for visuospatial processing and motor planning¹¹. The involvement of motor planning was indicated by substantial activation in the premotor cortex (BA6)^{9,10,12}, which is not active for purely visual images (as in mental rotation of letters or figures)^{12,13}. Compared with the 'old' instruction, the 'new' instruction resulted in much more activation in the premotor, parietal and prefrontal cortices (Fig. 4c, d). Activation in the inferior part of the posterior frontal cortex (BA44), intraparietal sulcus (IPS; BA7/39) and prefrontal cortex (BA47, 46, 9) were unique to the 'new' instruction (Fig. 4c, d). The activity in the inferior posterior frontal cortex was limited to the left hemisphere regardless of whether the imagined hand was the left or right hand. From the sulcal pattern, the region of activity was likely to be in Broca's area (BA44; Fig. 4e–h). These results indicate

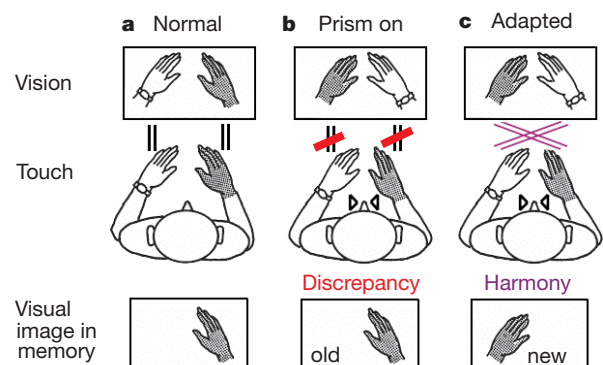


Figure 3 Schematic description of visuomotor experience for a subject's own right hand, before and while wearing the prisms. In our definition, the hand representation stores contingency between motor commands and visual images of hand. **a**, Before wearing prisms. Because normal contingency is stored (parallel lines), the visual image of the right hand is the same as the visible right hand. **b**, Wearing the prisms. The normal (old) contingency leads to an expected view (visual image) of the right hand that is the reverse of the visible hand for the prism wearer. Thus the subject experiences discrepancy. **c**, After adaptation, the new contingency (purple lines) is stored. This leads to a new visual image that is identical to the visible one. Thus the subject perceives harmony.

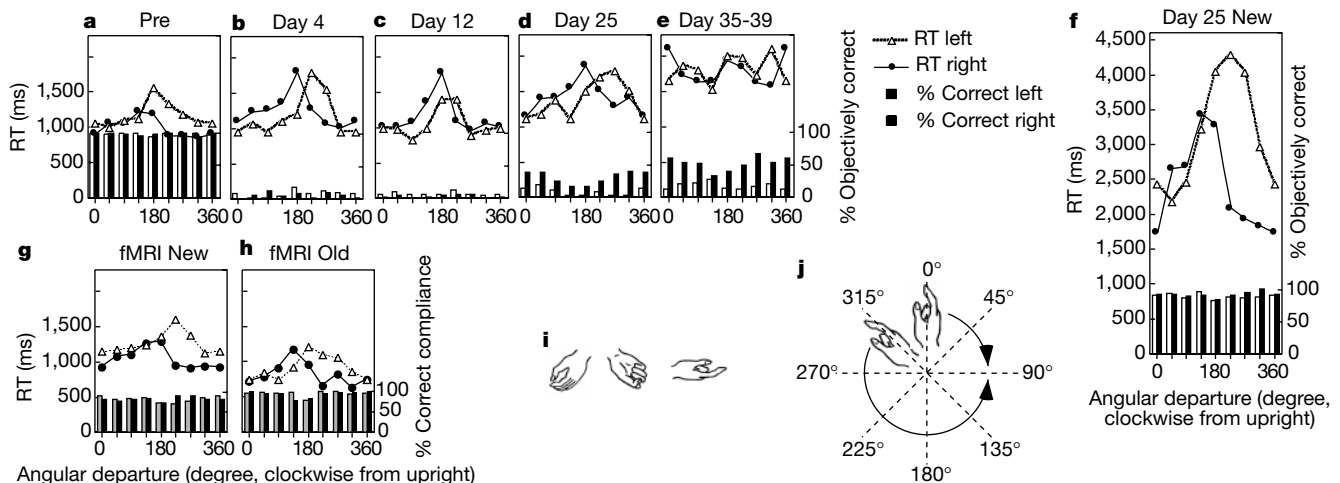


Figure 2 The hand identification task. **a–f**, Reaction time (RT) experiment. Mean RT and per cent correct identification are shown as a function of orientation of the stimuli, for the pre-test (**a**), the ordinary instruction in adaptation period (**b–e**), and the 'new-image' instruction (**f**). **g,h**, Performance during fMRI scans (days 34–38, see Methods). Owing to inability to respond within the 2.6-s time limit, 12% of the trials for the 'new' instruction,

and 3% for the 'old', resulted in missing RT. Analyses of variance (ANOVAs) showed that the 'new' instruction tended to yield longer RT ($P < 0.075$) and fewer correct responses ($P < 0.075$) than did the 'old'. **i**, Examples of stimuli. **j**, Eight stimulus orientations and two prototypical orientations (0° and 315°) for the right hand.

that accessing the new hand image is mediated by parieto-frontal circuits including the IPS, Broca's area and the prefrontal cortex^{14,15}.

Presumably the hand representation (memory storage of visuomotor mapping, in our definition) *per se* cannot be localized in certain regions in the brain. Rather, we may only observe the representation as widespread connections when it is active. Supporting this view, there are anatomically identified parieto-frontal circuits in monkey in which BA7 (around the IPS) projects to ventral premotor cortex (PMv) and prefrontal cortex (PF), and PF projects to the PMv^{14,15}. The PF often subserves associative learning^{16,17} or working memory^{18,19}, and BA7 and PMv are predominantly engaged in visuomotor transformation^{20–24}. The observed activity in the PF (here, BA47, 46, 9) would reflect greater resource allocation for recalling the new hand image, presumably because the visuomotor association in the new mapping is weaker than that in the old. The greater PF activation for the left-hand image than for the right (Fig. 4c, d) is consistent with the behavioural data, indicating slower acquisition of the new image for the left hand (Fig. 2d–f).

The visuomotor function of BA7 has been well documented in both human^{20,21} and monkey^{22–24}, and this area may code the schema of hand²⁵. In contrast, Broca's area in the human brain, widely known for its language function, is less recognized for its visuomotor contribution. However, there is evidence that the human Broca's area is also involved in preparation for imitative hand movements^{26,27}. The observed activity in Broca's area presumably reflects a visuomotor transformation process that is essential for accessing the new hand representation. Consistent with our results,

the monkey PMv (part of which is homologous with human Broca's area, according to some researchers²⁸) has been shown to be crucial for prism adaptation²⁹. Also note that human BA7 is related to prism adaptation²¹. Both areas were active in our 'new' instruction, indicating that the new visuomotor mapping may be coded as a network encompassing many areas¹⁴.

In our psychophysical data, the visual localization and hand identification tasks showed a similar time course of adaptation, including slower adaptation for the left (Figs 1b, 2). The new hand representation would serve as a frame of reference upon which normal space perception is regained. According to introspective reports, perceptual harmony between the visible and tactile worlds may be restored only when the subject actively intends to reference the visible world to his or her own body^{1,2}. We infer that the new hand representation contains reversed visuomotor mapping (Fig. 3c), and operates as a visuomotor transformation device that transforms visual images (of objects) into reversed motor output. If the reversed mapping operates on visual input coming through the prism, it would produce normal motor output and would therefore virtually cancel out the reversal of visual input. Although visually guided behaviour measured in the reaching task was restored within 2 weeks (Fig. 1a), more complex behaviour such as riding a bicycle was possible only after the emergence of the new hand representation. This may indicate that the hand representation predicts the relationship between visual and motor codes in various contexts.

Space coding would be often arm-centred⁶ in the normal situation as well as in adaptation to reversed vision, because we recognize the visual world primarily to allow active interaction between objects and our hands. □

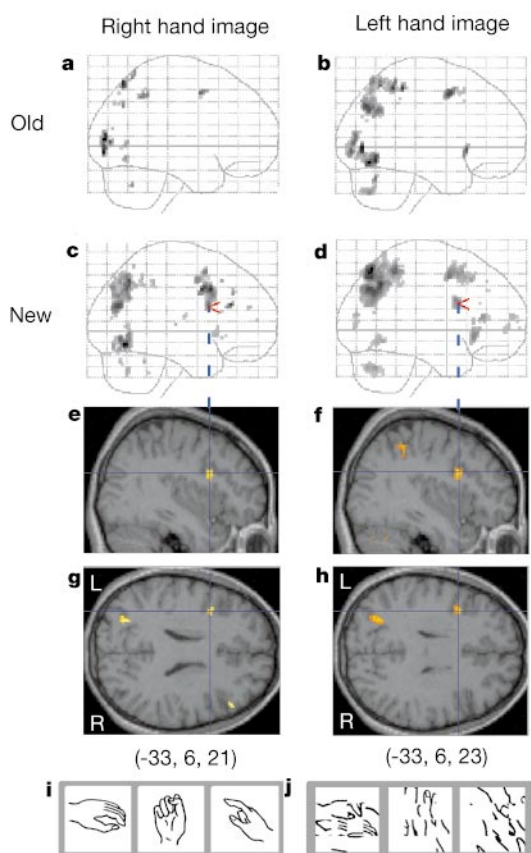


Figure 4 Functional MRI scans during mental rotation of a subject's own hand image. Significantly activated voxels ($P < 0.05$, corrected) are shown as sagittal projections for each condition (a–d). The activity in BA44 is also shown in section displays (e–h) with Talairach coordinates for the maxima (see also Supplementary Information) (indicated with red markers in c and d). Example stimuli are shown for the test (i) and control periods (j).

Methods

Subjects were tested with the reversing spectacles on except for the pre- and post-tests, for which they wore only the frame of the spectacles. All subjects gave informed consent before the experiment.

Reaching task

The subject's head was placed on a chin rest at a viewing distance of 45 cm so that the tested eye (right eye for JA, TU and YH; left eye for CH) was located in front of the centre of a touch display. The other eye was covered. A trial began with a fixation point at the centre of the display, which disappeared when the subject pressed and held a key. After a variable period, a visual target (a circle, 16 mm in diameter) was presented for 500 ms at one of nine locations: the centre and eight radial locations (81 mm from the centre). The subjects were instructed to reach the target as fast and accurately as possible using the preferred (right) hand. The latency (for the hand to depart from the ready position), movement time and the coordinates of the location reached were recorded.

Visual localization task

The subject sat in front of a display with a viewing distance of 105 cm. He/she placed both hands on a switch box that was hidden from view on the subject's knees. A trial began with a fixation point at the centre of the display being displayed for 500 ms. After a 100-ms period with no display, a visual target (a cross, 3 cm × 3 cm) was presented for 200 ms at either 10.5 cm left or right of the centre. The subjects were instructed to press the button that they perceived to be on the same side (left or right) as the visual target, as fast as possible.

Hand identification task (RT experiment)

The subject's head was placed on a chin rest at a viewing distance of 45 cm. Their hands were kept stationary and hidden on their knees under a table. Stimuli consisted of three line drawings of a hand (Fig. 2i) that were presented as either a left or right hand in one of eight orientations in the picture plane (Fig. 2j). When the subject was ready, a series of the total 48 stimuli began in random order. The subjects were instructed to identify each hand as left or right and report the handedness verbally. A stimulus stayed visible until the onset of the subject's voice. During prism wearing, the ordinary instruction asked the subjects to select the more vivid image, in case they had both old and new hand images. In the new-image condition on day 25, they were instructed to exclusively attend to the new hand image.

fMRI measurement

Echo planar MRI data were acquired with a 1.5-T Siemens Vision system with a standard CP head coil. T_2^* -weighted functional images were acquired with an in-plane resolution of 2.34 mm × 2.34 mm (repetition time 5.3 s; echo time 66 ms) at each of 16 axial

noncontiguous 7-mm-thick slices (with 3.5-mm interslice gaps). In each functional sequence, eight test and eight control scans alternated for a total of 120 scans. In the test periods, the above hand identification task was given, with one trial every 2.6 s (Fig. 4i). The subjects were asked to report the handedness of each stimulus within 2.6 s by pressing either a head-side or foot-side key on the abdomen with the index fingers of both hands simultaneously. The stimulus disappeared upon response or at the end of the 2.6 s. In the control periods, the hand stimuli were replaced by scrambled hands (Fig. 4j). The subjects were asked to report whether the majority of the lines were near oblique or horizontal/vertical by key pressing. In the left-hand sequence, 81% (13/16) of the stimuli were left hands, and vice versa in the right-hand sequence. The subjects were tested under both new and old image instructions for both left and right hand sequences on days 34–38. A post-test was conducted five months after removal of the reversing spectacles. The post-test was equivalent to the old-image instruction because the subjects claimed that the new hand image did not make sense at that time.

All fMRI data were processed using the SPM99 software package (<http://www.fil.ion.ucl.ac.uk>). Standard linear image realignment, linear normalization to the stereotactic anatomical space and spatial smoothing (three-dimensional gaussian kernel, 4.7 mm full-width at half-maximum) were successively performed for each subject. Then, the data from subjects were pooled together and group comparisons were performed using the general linear model. Comparing hand identification with orientation judgement, significant increases were tested with *t* statistics and displayed as statistical parametric maps. Threshold for significance was set at voxel-level $P < 0.05$, corrected for multiple comparisons.

Received 22 May; accepted 14 July 2000.

- Kohler, I. The formation and transformation of the perceptual world. *Psychol. Issues* 3 (monogr. 12) 1–173 (1964).
- Stratton, G. M. Vision without inversion of the retinal image. *Psychol. Rev.* 4, 341–360; 463–481 (1897).
- Sugita, Y. Global plasticity in adult visual cortex following reversal of visual input. *Nature* 380, 523–526 (1996).
- Miyachi, S. et al. Adaptation to reversing prisms activates the ipsilateral visual cortex in humans: an fMRI study. *Invest. Ophthalmol. Vis. Sci.* 40, S820 (1999).
- Held, R. Shifts in binocular localization after prolonged exposure to atypical combination of stimuli. *Am. J. Psychol.* 68, 526–548 (1955).
- Graziano, M. S. A., Yap, G. S. & Gross, C. G. Coding of visual space by premotor neurons. *Science* 266, 1054–1057 (1994).
- Sekiyama, K. Kinesthetic aspects of mental representations in the identification of left and right hands. *Percept. Psychophys.* 32, 89–95 (1982).
- Sekiyama, K. Mental and physical movements of hands: Kinesthetic information preserved in representational systems. *Jpn. Psychol. Res.* 25, 95–102 (1983).
- Bonda, E., Petrides, M., Frey, S. & Evans, A. Neural correlates of mental transformations of the body-in-space. *Proc. Natl Acad. Sci. USA* 92, 11180–11184 (1995).
- Parsons, L. M. et al. Use of implicit motor imagery for visual shape discrimination as revealed by PET. *Nature* 375, 54–58 (1995).
- Roland, P. E. *Brain Activation* (Wiley-Liss, New York, 1993).
- Kosslyn, S. M., DiGirolamo, G. J., Thompson, W. L. & Alpert, N. M. Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. *Psychophysiology* 35, 151–161 (1998).
- Alivisatos, B. & Petrides, M. Functional activation of the human brain during mental rotation. *Neuropsychologia* 35, 111–1118 (1997).
- Jeannerod, M., Arbib, M. A., Rizzolatti, G. & Sakata, H. Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends Neurosci.* 18, 314–320 (1995).
- Selemon, L. D. & Goldman-Rakic, P. S. Common cortical subcortical targets of the dorsolateral prefrontal and posterior parietal cortices in the rhesus monkey: Evidence for a distributed neural network subserving spatially guided behavior. *J. Neurosci.* 8, 4049–4068 (1988).
- Murray, E. A., Bussey, T. J. & Wise, S. P. Role of prefrontal cortex in a network for arbitrary visuomotor mapping. *Exp. Brain Res.* 1, 114–129 (2000).
- White, I. M. & Wise, S. P. Rule-dependent neuronal activity in the prefrontal cortex. *Exp. Brain Res.* 126, 315–335 (1999).
- Cohen, J. D. et al. Temporal dynamics of brain activation during a working memory task. *Nature* 386, 604–608 (1997).
- Courtney, S. M., Ungerleider, L. G., Keil, K. & Haxby, J. V. Transient and sustained activity in a distributed neural system for human working memory. *Nature* 386, 608–611 (1997).
- Perenin, M. T. & Vighetto, A. Optic ataxia: A specific disruption in visuomotor mechanisms. *Brain* 111, 643–674 (1988).
- Clover, D. M. et al. Role of posterior parietal cortex in the recalibration of visually guided reaching. *Nature* 383, 618–621 (1996).
- Snyder, L. H., Batista, A. P. & Andersen, R. A. Coding of intention in the posterior parietal cortex. *Nature* 386, 167–170 (1997).
- Rushworth, M. F. S., Nixon, P. D. & Passingham, R. E. Parietal cortex and movement: I. Movement selection and reaching. *Exp. Brain Res.* 117, 292–310 (1997).
- Taira, M., Mine, S., Georgopoulos, A. P., Murata, A. & Sakata, H. Parietal cortex neurons of the monkey related to the visual guidance of hand movement. *Exp. Brain Res.* 83, 29–36 (1990).
- Iriki, A., Tanaka, M. & Iwamura, Y. Coding of modified body schema during tool use by macaque postcentral neurons. *NeuroReport* 7, 2325–2330 (1996).
- Krams, M., Rushworth, M. F. S., Deiber, M. P., Frackowiak, R. S. J. & Passingham, R. E. The preparation, execution and suppression of copied movements in the human brain. *Exp. Brain Res.* 120, 386–398 (1998).
- Grafton, S. T., Arbib, M. A., Fadiga, L. & Rizzolatti, G. Localization of grasp representations in humans by positron emission tomography: 2. Observation compared with imagination. *Exp. Brain Res.* 112, 103–111 (1996).
- Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141 (1996).

29. Kurata, K. & Hoshi, E. Reacquisition deficits in prism adaptation after muscimol microinjection into the ventral premotor cortex of monkeys. *J. Neurophysiol.* 81, 1927–1938 (1999).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank M. Kato, T. Hayakawa, T. Murata, H. Tanabe and M. Nakatsuka for technical assistance and discussions.

Correspondence and requests for materials should be addressed to K.S. (e-mail: sekiyama@fun.ac.jp).

IRS-2 pathways integrate female reproduction and energy homeostasis

Deborah J. Burks*, Jaime Font de Mora†, Markus Schubert*, Dominic J. Withers‡, Martin G. Myers*, Heather H. Towery*, Shari L. Altamuro*, Carrie L. Flint* & Morris F. White*

* Howard Hughes Medical Institute, Joslin Diabetes Center, Harvard Medical School, One Joslin Place, Boston, Massachusetts 02215, USA

† Centro de Investigacion del Cancer, Facultad de Medicina, Universidad de Salamanca, Salamanca 37007, Spain

Severe dietary restriction, catabolic states and even short-term caloric deprivation impair fertility in mammals. Likewise, obesity is associated with infertile conditions such as polycystic ovary syndrome^{1,2}. The reproductive status of lower organisms such as *Caenorhabditis elegans* is also modulated by availability of nutrients^{3,4}. Thus, fertility requires the integration of reproductive and metabolic signals. Here we show that deletion of insulin receptor substrate-2 (IRS-2), a component of the insulin/insulin-like growth factor-1 signalling cascade, causes female infertility. Mice lacking IRS-2 have small, anovulatory ovaries with reduced numbers of follicles. Plasma concentrations of luteinizing hormone, prolactin and sex steroids are low in these animals. Pituitaries are decreased in size and contain reduced numbers of gonadotrophs. Females lacking IRS-2 have increased food intake and obesity, despite elevated levels of leptin. Our findings indicate that insulin, together with leptin and other neuropeptides, may modulate hypothalamic control of appetite and reproductive endocrinology. Coupled with findings on the role of insulin-signalling pathways in the regulation of fertility, metabolism and longevity in *C. elegans* and *Drosophila*^{3–5}, we have identified an evolutionarily conserved mechanism in mammals that regulates both reproduction and energy homeostasis.

Insulin receptor substrate (IRS) proteins undergo rapid tyrosine phosphorylation in response to insulin and insulin-like growth factor-1 (IGF-1). The mammalian IRS protein family contains at least four members: IRS-1 and IRS-2, which are widely expressed; IRS-3, which is found predominantly in adipose tissue; and IRS-4, which is expressed in the thymus, brain and kidney^{6,7}. We have shown, using murine gene deletion, that IRS-2 is critical for peripheral carbohydrate metabolism and β -cell function^{8,9}. In addition, IRS-1 mediates embryonic and post-natal somatic growth⁸. In *C. elegans*, the insulin/IGF-1 receptor homologue DAF-2 through the AGE-1 phosphatidylinositol-3-OH kinase (PI(3)K), regulates development, reproduction and longevity in response to environmental signals such as food³. Mutations in these pathways can induce developmental arrest at the dauer stage, reduce

‡ Present address: Department of Metabolic Medicine, Imperial College School of Medicine, Hammer-smith Campus, DuCane Road, London W12 0NN, UK.

fertility and/or extend life-span^{4,10}. Similarly, deletion of CHICO, the IRS protein of *Drosophila*, causes female sterility and reduced somatic growth and increased lipid storage⁵.

During our analysis of the role of IRS-2 in carbohydrate metabolism⁸, we found that matings between IRS-2^{-/-} females and IRS-2^{-/-} males did not yield a single pregnancy. To characterize the impaired reproduction, we established controlled breedings using various IRS genotypes (Fig. 1a). Only 9% of IRS-2^{-/-} females (4–6-week-old virgins) that mated with wild-type males became pregnant during an 8-week period, and this was reduced to 0% when IRS-2^{-/-} females were bred with IRS-2^{-/-} males (Fig. 1a). IRS-2^{-/-} males also displayed reduced fertility. If mated before the onset of severe diabetes, however, IRS-2^{-/-} males were adequate breeders. Detailed analysis of our knockout colony revealed a sexual dimorphism in the diabetic phenotype: defects in carbohydrate metabolism imposed by deletion of IRS-2 were less severe in females than in males. Although IRS-2^{-/-} males were severely glucose intolerant by 6 weeks, IRS-2^{-/-} females displayed only slightly elevated blood glucose levels and mildly impaired glucose tolerance (Fig. 1b, c). IRS-2^{-/-} females maintained fasting glucose levels in the range 120–160 mg dl⁻¹ until about 4–5 months of age. All females used in the reproductive study were less than 10 weeks of age and so were relatively euglycaemic and mildly insulin resistant. It is therefore unlikely that the profound disturbance in fertility is a direct consequence of abnormal glucose metabolism in female IRS-2^{-/-} animals.

We monitored IRS-2 matings daily for the presence of vaginal plugs—an indication of copulation and indirect evidence for lordotic behaviour. Copulation plugs were rarely noted in IRS-2^{-/-} females during the course of this study (18.2% of IRS-2^{-/-} compared with 100% of wild type). These observations suggest dysregulation of the oestrous cycle and normal sexual behaviour rather than failed implantation as the underlying cause of the infertility. Gross anatomical examination of 6-week-old IRS-2^{-/-} mice revealed normal development of the external genitalia and the reproductive tract; however, the ovaries of IRS-2^{-/-} mice were small, contained very few surface follicles and displayed thickening of the cortex (Fig. 2a). Histological examination of ovarian sections from

IRS-2^{-/-} females revealed further evidence of anovulation including thickening of the ovarian stroma and an almost complete absence of corpora lutea (number of corpora lutea/ovary in wild type, 4.2 ± 0.6, *n* = 9; in IRS-2^{-/-}, 0.4 ± 0.2, *n* = 12) (Fig. 2b, c). Adult ovaries also contained decreased numbers of primary follicles with few growing follicles reaching an antral phase of development (Fig. 2c).

Quantitation of primary oocytes from embryos at day 18.5 revealed reduced numbers of these cells in the IRS-2^{-/-} ovary as compared with wild-type controls (Fig. 2d), suggesting that the absence of IRS-2 impairs proliferation and/or increases apoptosis in this cell population during ovarian development. Notably, IGF-I and activation of PI(3)K have been identified as crucial mediators of germ-cell survival during murine embryonic oogenesis¹¹. IRS-2 may therefore represent an important component of IGF-I and/or other developmental signalling pathways during fetal oogenesis. Both IRS-1 and IRS-2 were present in normal adult ovaries and phosphorylated during insulin stimulation (Fig. 2e). Expression and insulin-stimulated phosphorylation of IRS-1 in IRS-2^{-/-} ovaries were comparable to levels detected in control animals (Fig. 2e), suggesting either that IRS-2 pathways transmit a specific function in this tissue or that normal development and endocrine regulation of ovarian physiology requires a signalling balance between IRS molecules.

To investigate further the defective reproduction in IRS-2^{-/-} females, we measured sex steroid hormones and luteinizing hormone levels during stages of the oestrous cycle. Many IRS-2 knockouts (61%) failed to cycle but remained permanently in an inactive or dioestrous state. In both dioestrous and oestrous states, sex steroid hormones were reduced in IRS-2 knockouts (Fig. 3a). Normally, if ovarian feedback signals to the pituitary–hypothalamic axis are removed, as in the case of ovarian failure or ovariectomy, luteinizing hormone levels are elevated owing to loss of sex steroid suppression of gonadotropin-releasing hormone. Unexpectedly, luteinizing hormone levels were low in IRS-2^{-/-} animals, suggesting that there are defects in production and/or release of this hormone. Elevated prolactin is commonly associated with human infertility and we therefore assayed circulating prolactin levels of IRS-2^{-/-} female mice. Consistent with observations regarding other reproductive

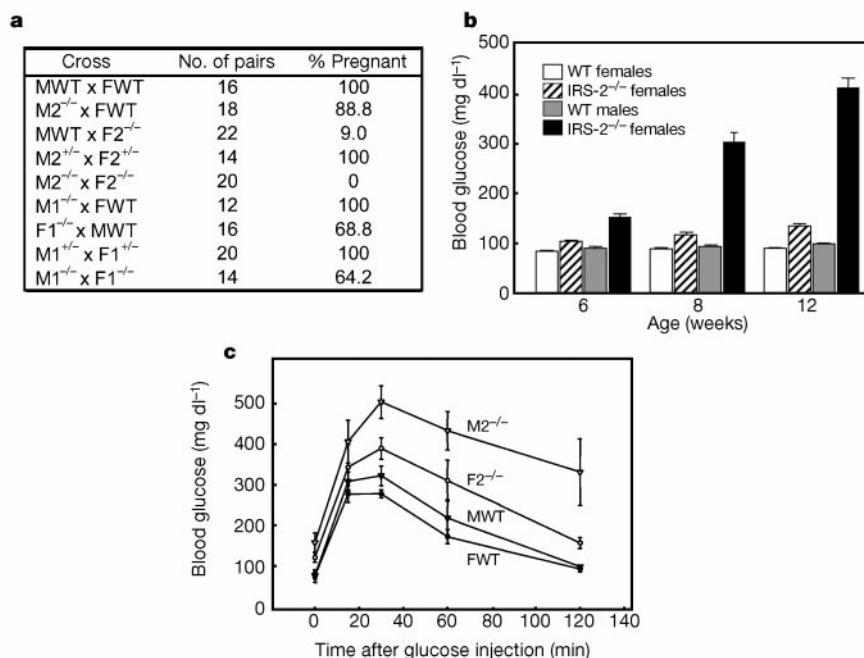


Figure 1 Reproductive and metabolic characteristics of IRS-2-deficient animals.

a, Percentage of pregnancies. 4–6-week-old females were mated with experienced males of the indicated genotypes. Matings were monitored for pregnancies during an 8-week period. MWT, male wild-type; FWT, female wild-type; M2^{-/-}, male IRS-2 knockout; F2^{-/-},

female IRS-2 knockout; M1^{-/-}, male IRS-1 knockout; F1^{-/-}, female IRS-1 knockout. **b**, Blood glucose levels in 6-, 8-, and 12-week-old mice after a 15-h overnight fast. **c**, Glucose tolerance tests in fasted 6-week-old mice after intraperitoneal injection of 2 g o-glucose per kg body weight. Results in **b** and **c** are the mean ± s.e.m. of six mice of each genotype.

hormones, prolactin levels were reduced in these animals (wild-type, $10.6 \pm 2.2 \text{ ng ml}^{-1}$, $n = 5$ mice; IRS-2^{-/-}, $6.6 \pm 1.2 \text{ ng ml}^{-1}$, $n = 8$ mice). IRS-2^{-/-} females were also resistant to exogenous gonadotropin stimulation: we retrieved only 0.25 ± 0.01 oocytes per animal ($n = 8$ animals) from the oviducts of superovulated IRS-2^{-/-} mice, but 12 ± 0.8 oocytes per animal in wild-type controls ($n = 6$ animals).

The gene for IRS-2 has been implicated as being regulated by gonadal steroids¹². As endogenous levels of oestrogen and progesterone are reduced in IRS-2 knockouts, we evaluated their response to exogenous sex steroids. Ovariectomized 6-week-old mice were treated for 3 weeks with a daily dose of 1 mg oestradiol or 1 µg progesterone, or a combination of both hormones. Uterine weight and morphology of treated knockouts were comparable to control values under all three experimental conditions, suggesting that uterine proliferation in IRS-2^{-/-} females proceeded normally when regulated by exogenous steroids (Fig. 3b; data not shown). Oestradiol stimulates tyrosine phosphorylation of IRS-1 in uterine epithelial cells¹³. Short-term hormonal treatment (24 h) enhanced phosphorylation of IRS-1 in uterine extracts from IRS-2^{-/-} females and controls (Fig. 3c). Moreover, hormonal induction of oestrogen receptor expression in IRS-2^{-/-} tissue was similar to the levels detected in wild-type animals (Fig. 4c). These experiments show that tissues of the IRS-2 knockout respond appropriately to exogenous sex steroids.

To explore the potential defects of the reproductive axis that contributed to the low hormone levels, we analysed pituitary morphology in IRS-2^{-/-} females. Pituitaries were reduced in size

by roughly 30% (pituitary weight: wild-type, $45.6 \pm 1.2 \text{ mg}$, $n = 9$; IRS-2^{-/-}, $31.4 \pm 2.1 \text{ mg}$, $n = 12$) (Fig. 4a). Although pituitary size was globally diminished, disruption of IRS-2 particularly reduced the intermediate lobe, as visualized by immunostaining with anti-ACTH (adrenocorticotrophic hormone) antibodies (Fig. 4a, b). To assess the effect of pituitary size reduction on endocrine function, we immunostained pituitary sections for growth hormone and the gonadotropins luteinizing hormone and follicle-stimulating hormone (FSH). The number of gonadotrophs was reduced by almost 40% in the IRS-2^{-/-} pituitaries, whereas somatotrophs were equivalent between knockouts and wild-type animals (Fig. 4c). Reduction of the gonadotroph population implicates IRS-2 pathways in the development and maintenance of these hormone-producing cells and provides at least one explanation for the abnormal levels of luteinizing hormone and FSH in the IRS-2^{-/-} females.

Many studies suggest that insulin serves as a link between the periphery and the brain at various levels of the neuroendocrine axis¹⁴. The receptors for insulin and IGF-1 show distinct patterns of expression within the pituitary gland; these receptors are detected on different subpopulations of secretory cells of the pars distalis^{15,16}. Moreover, receptors for insulin are expressed by a subset of cells within the intermediate lobe¹⁵. To examine the capacity of IRS-2 to mediate insulin/IGF-1 signals in the neuroendocrine axis, we collected pituitaries after *in vivo* insulin stimulation of wild-type mice. Lysates of these glands were immunoprecipitated with either anti-IRS-1 or anti-IRS-2 antibodies and subsequently probed with anti-phosphotyrosine antibodies. Notably, there was tyrosine

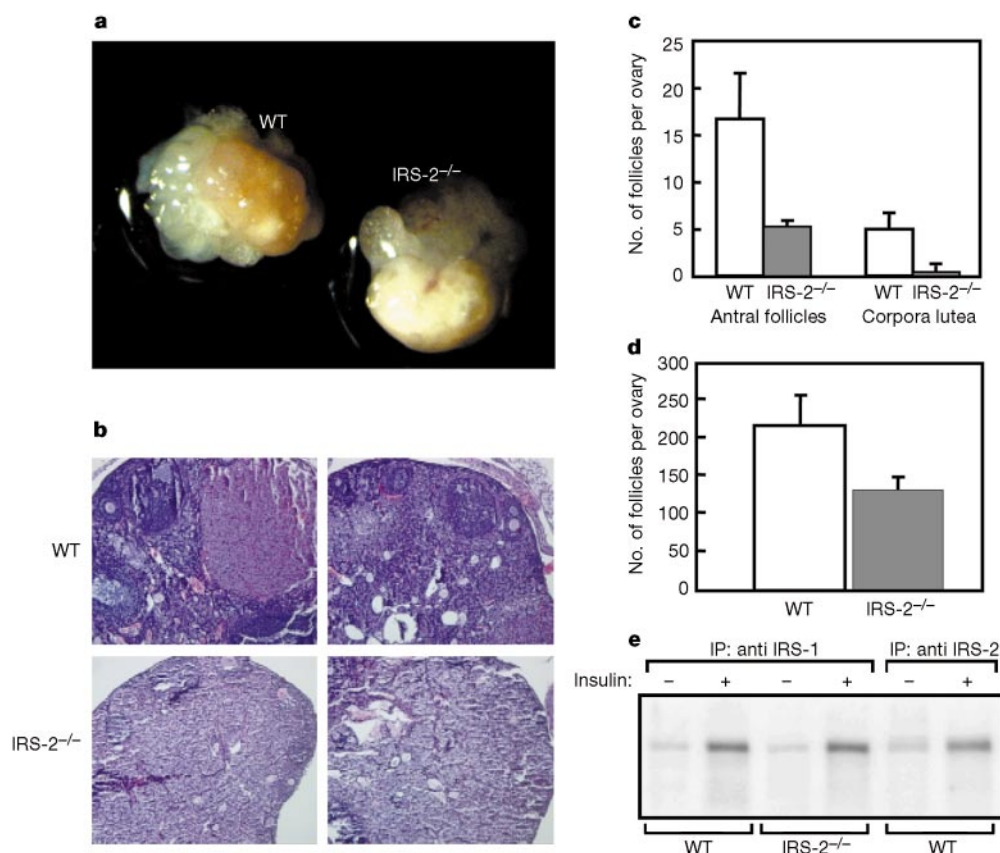


Figure 2 Ovarian phenotype of the IRS-2 knockout. **a**, Ovaries with attached oviducts and fallopian tubes from 8-week-old mice at $\times 2$ magnification. **b**, Ovaries from 6–8-week-old mice. Sections were stained with haematoxylin and eosin and viewed at $\times 20$ magnification. Representative ovarian sections are from two different animals of each genotype. **c**, Antral follicles (200–250 µm) and corpora lutea quantified from haematoxylin and eosin stained ovarian sections from 6–8-week-old mice viewed at $\times 20$ magnification. Results are the mean $\pm$ s.e.m. of at least eight sets of ovaries from each genotype. **d**, Quantification of primary germ cells at embryo day 18. Fetal ovaries were

collected and fixed in 4% paraformaldehyde. Primary oocytes were quantitated from haematoxylin and eosin stained sections of embryonic ovaries. Results are the mean $\pm$ s.e.m. of three females per genotype. **e**, Immunoblot of ovarian lysates. After *in vivo* insulin stimulation (5 units human insulin), ovaries were collected and homogenized. Lysates (1 mg total protein) were immunoprecipitated with anti-IRS antibodies as indicated, and immunocomplexes probed with anti-phosphotyrosine antibodies. Results are representative of three independent experiments.

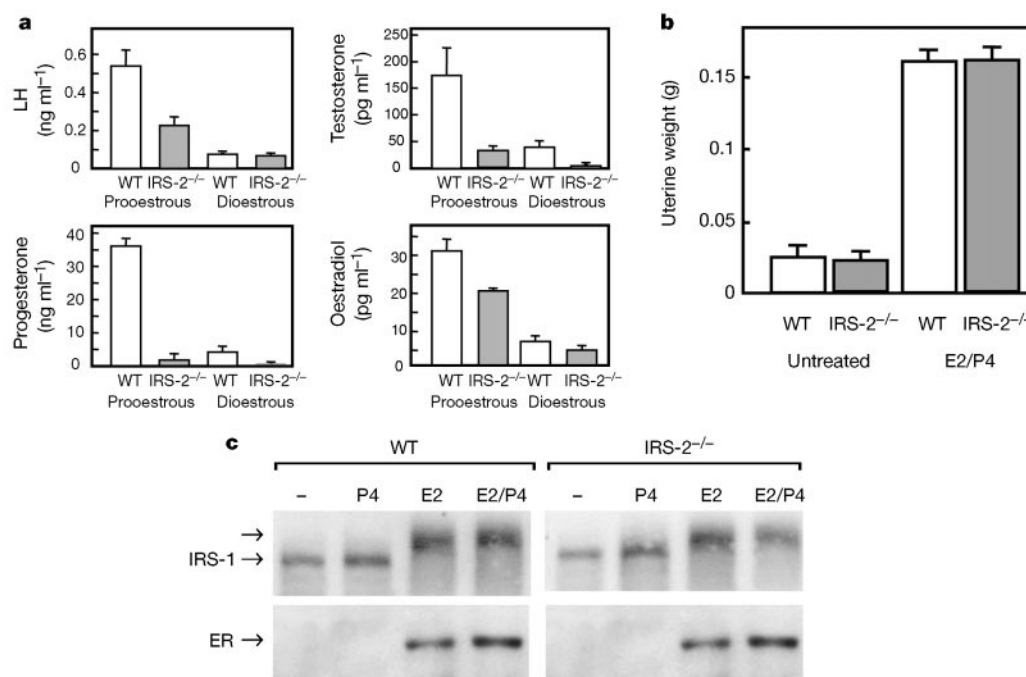


Figure 3 Reproductive hormone levels and response of ovariectomized mice to exogenous sex steroid stimulation. **a**, Plasma levels of reproductive hormones. Daily vaginal smears were used to assess the stage of the oestrous cycle. Serum was obtained from either pro-oestrous or dioestrous 8-week-old females by retro-orbital bleed. Wild-type pro-oestrous, $n = 6$; wild-type dioestrous, $n = 8$; IRS-2^{-/-} pro-oestrous, $n = 6$; IRS-2^{-/-} dioestrous, $n = 15$. **b**, Uterine weight. Ovaries were surgically removed from 6-week-old mice. Mice were allowed to recover for 2 weeks and then given daily subcutaneous doses of 1 mg oestradiol (E2) or 1 μ g progesterone (P4), or both. Control

animals were given vehicle mineral oil. After 21 days mice were killed, and uteri were dissected, cleared of fat and weighed to evaluate their proliferative response to sex steroid stimulation. Untreated animals, $n = 4$ females of each genotype; E2/P4 treated animals, $n = 6$ animals of each genotype. **c**, Immunoblots of uterine extracts. Lysates were immunoprecipitated for IRS-1, and immunoprecipitates were probed with anti-IRS-1 antibodies. E2 treatment reduced the mobility of IRS-1, indicative of enhanced phosphorylation. Total lysates (100 μ g protein) were also probed for oestradiol-mediated induction of oestrogen receptor expression.

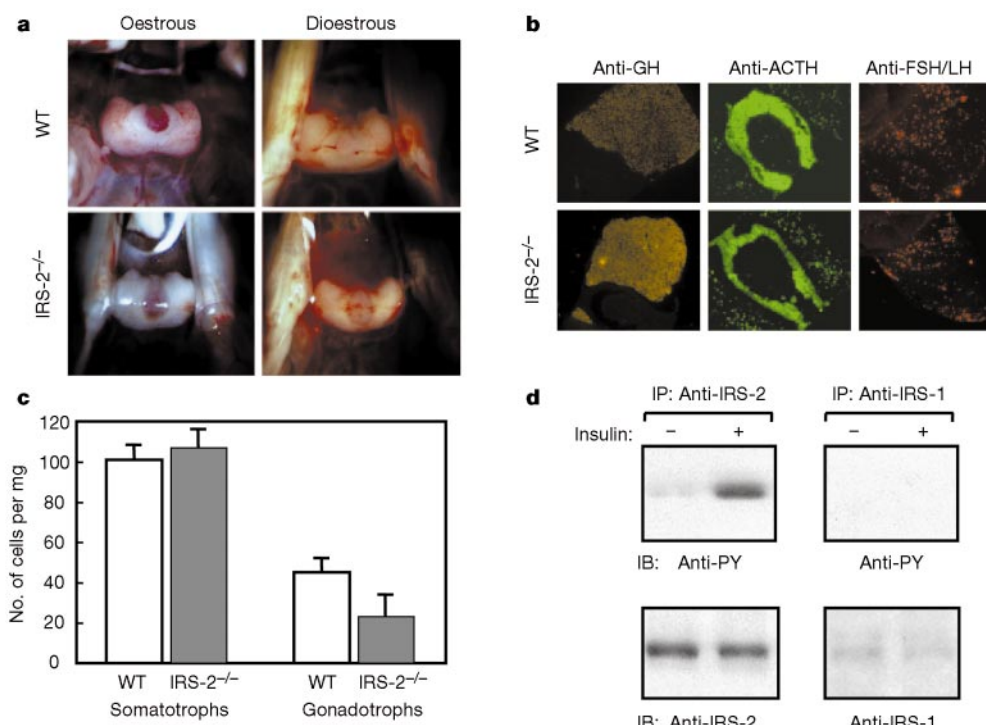


Figure 4 Pituitary phenotype of IRS-2^{-/-} females. **a**, Representative pituitaries from oestrous and dioestrous 6–8-week-old females of each genotype at $\times 2$ magnification. **b**, Immunofluorescent images of pituitary sections at $\times 20$ magnification. Sections (1 μ m) were co-immunostained for antibodies to growth hormone (GH), adrenocorticotrophic hormone (ACTH), follicle stimulating hormone (FSH) and luteinizing hormone (LH). **c**, Quantitation of somatotrophs and gonadotrophs. Pituitary sections from 6–8-week-old females were co-immunostained for growth hormone, ACTH and FSH/LH. Results are

mean $\pm$ s.e.m. of four mice from each genotype. **d**, Immunoblot of wild-type pituitary lysates after acute *in vivo* insulin stimulation (5 units, 10 min). Lysates (0.5 mg total protein) were immunoprecipitated with anti-IRS antibodies as indicated, and immunoprecipitates were probed with anti-phosphotyrosine antibodies. Pituitary lysates were pooled from three untreated or three insulin-stimulated mice. Results are representative of two independent experiments.

phosphorylation of IRS-2, but not of IRS-1 (Fig. 4d). Both molecules were present and tyrosine phosphorylated in liver lysates from these stimulated animals (data not shown). These results therefore suggest that IRS-2 pathways, based on either a distinct pattern of expression or a unique signalling capacity, may regulate normal development and function of the pituitary.

In *C. elegans* and *Drosophila*, homologues of the insulin/IGF-1/IRS signalling network have been implicated not only in the regulation of reproduction but also in the maintenance of fuel homeostasis^{3–5}. In addition, insulin acts centrally to regulate appetite in mammals^{14,17,18}. We therefore looked for evidence that IRS-2-dependent mechanisms might co-ordinate reproduction with energy balance. Monitoring of food intake showed that IRS-2^{-/-} mice consumed 30% more chow than control animals (IRS-2^{-/-}, 4.5 ± 0.30 g day⁻¹ per mouse; wild-type, 3.1 ± 0.19 g day⁻¹ per mouse) (Fig. 5a). Furthermore, this increased food intake was reflected by abnormalities in fuel storage; IRS-2^{-/-} mice weighed 20% more and stored two times more body fat than age-matched controls (Fig. 5b, c).

Circulating leptin levels were elevated more than 5-fold in the 8-week-old IRS-2^{-/-} females as compared with controls, consistent with their increased adiposity (Fig. 5d). Moreover, leptin levels in IRS-2^{-/-} females were elevated 2.5-fold (IRS-2^{-/-}, 6.1 ± 0.8 ng ml⁻¹; wild-type, 2.4 ± 0.4 ng ml⁻¹) as early as 4 weeks of age, before the onset of abnormal glucose tolerance. Leptin transmits critical signals regarding weight regulation to the central nervous system^{14,19–21}. Genetic leptin resistance resulting from mutations in the murine leptin receptor (*db/db* mice) is associated with hyperphagia, weight gain and female infertility^{22–24}. The elevated leptin levels and feeding abnormalities implicate leptin resistance as one explanation for the dysregulated energy homeostasis in IRS-2^{-/-} females. We therefore assayed STAT3 phosphorylation in the hypothalamus of mice treated with an acute intravenous dose of leptin (0.5 µg g⁻¹ body weight). Leptin receptor activation in the hypothalamus stimulates phosphorylation and nuclear translocation of STAT3; dysregulation of this pathway is associated with

leptin resistance in *db/db* mice^{25,26}. Hypothalamic lysates were immunoprecipitated with anti-STAT3 antibodies and immunocomplexes were subsequently probed with phosphospecific-STAT3 antibodies. Leptin stimulated phosphorylation of STAT3 in wild-type hypothalami (Fig. 5e); however, leptin-treated IRS-2^{-/-} females displayed little or no activation of this pathway (Fig. 5e). Hypothalamic expression of STAT3 was equivalent between wild-type and knockout animals (Fig. 5e). These results show that IRS-2^{-/-} females have hypothalamic resistance to leptin, and suggest that in the absence of IRS-2 pathways homeostatic mechanisms required for leptin sensing and/or signalling are impaired. Leptin has been reported to activate some components of the insulin signalling including IRS-1 and IRS-2 (refs 27–29), suggesting functional interactions between the insulin and leptin pathways.

Our results implicate IRS-2 pathways in the coordination of reproduction and energy homeostasis. Female IRS-2^{-/-} mice have defects at more than one level of the hypothalamic–pituitary–ovarian axis, coupled with increased food intake and fat storage. IRS-2 may have evolved a unique role in this regulation, as IRS-1^{-/-} mice do not present severe abnormalities in reproductive physiology (Fig. 1a; unpublished data). In addition to its classical metabolic effects, insulin, as well as IGF-1, may modulate ovarian function¹. The IRS-2 model provides an important tool to unravel the complex pathogenesis of human infertility associated with insulin-resistant states such as polycystic ovarian syndrome². Dysregulated IRS-2 expression or function may represent one underlying molecular defect responsible for these human endocrine disorders.

The feeding abnormalities of IRS-2^{-/-} mice provide a molecular basis for the observation that intracerebroventricular administration of insulin decreases food intake in mammals^{14,17}. Peripheral insulin resistance caused by deletion of IRS-2 (ref. 8) is probably accompanied by a loss of sensitivity to insulin in critical brain regions. Consistent with this hypothesis, obesity in IRS-2^{-/-} mice occurs despite elevated leptin levels, suggesting that this hypothalamic leptin resistance is associated with both peripheral and central nervous system resistance to insulin action. Identification of

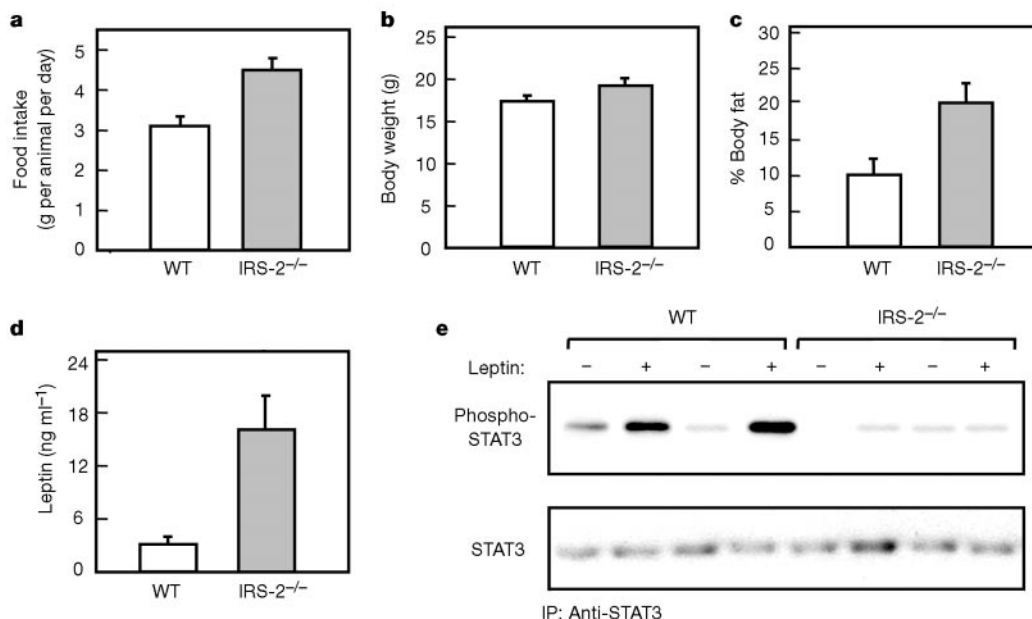


Figure 5 Analysis of feeding and energy homeostasis. **a**, Daily food intake. Female animals (6 weeks) were maintained on ground, normal mouse chow and monitored for 14 days. **b**, Body weight. Mice were weighed weekly from postnatal day 21 until 8 weeks of age. **c**, Body fat. Composition was determined by alkaline hydrolysis of carcasses and extraction of total lipids. **d**, Plasma leptin. Serum was collected by heart puncture and circulating leptin was measured by radioimmunoassay. Measurements were made on the same cohort of female mice in **a–d**; data are the mean ± s.e.m. of eight animals per

genotype. **e**, Immunoblot of hypothalamic lysates. Eight-week-old mice were fasted overnight for 14–16 h. Anesthetized mice were then injected intravenously with a bolus of leptin (0.5 µg g⁻¹ body weight). Untreated controls received an injection of saline. Two animals of each genotype were used for each condition. After 15 min, hypothalami were dissected and lysed. Lysates were immunoprecipitated with anti-STAT3 antibodies, and immunoprecipitates were then probed using phosphospecific STAT3 antibodies. Results represent two independent leptin stimulation experiments.

hypothalamic IRS-2 dependent signalling pathways will provide insights into the mechanisms of body weight regulation, and identify targets of therapeutic benefit. Our findings coupled with observations regarding insulin signalling pathways in *C. elegans* and *Drosophila* reveal an evolutionarily conserved mechanism by which neuroendocrine regulation of reproduction is coupled to appropriate energy metabolism. □

Methods

Mice

The generation of IRS-1 and IRS-2 knockout mice has been described^{8,9}. Both mouse lines were maintained on a mixed C57BL/6 × 129Sv genetic background. Genotyping of embryos and 3-week-old animals was done by Southern blotting as described⁸. Mice were maintained on normal light/dark cycle and handled in accordance with Joslin Diabetes Center Care and Use Committee protocols. Mating cages were monitored by 8:00 each day for the presence of vaginal plugs.

Hormone assays

To assess oestrous stages, we carried out daily vaginal smears. We collected blood for hormone assays by retro-orbital bleeds on anaesthetized mice or by tail bleeds. Murine luteinizing hormone and prolactin were measured by radioimmunoassay by T. Nett (Endocrine Laboratory of Colorado State University, Fort Collins). Sex steroids were assayed by the endocrine laboratory of the Brigham and Women's hospital and the laboratory of D. Hess (Oregon Regional Primate Center, Beaverton, OR). We determined leptin levels by radioimmunoassay using a mouse leptin standard (Linco).

Histology and immunostaining

We removed ovaries and uteri from adult animals and placed them in 10% buffered formalin. Tissues were embedded in paraffin, and 5-µm sections were prepared. Embryonic ovaries were collected in 4% paraformaldehyde. Pituitaries were fixed overnight in 4% paraformaldehyde and processed for electron microscopy. One-micron sections were stained with antibodies to growth hormone (Dako), and luteinizing hormone, FSH and ACTH, (A. F. Parlow, NIDDK, National Hormone and Pituitary Program). Detection was carried out using either rhodamine or fluorescein antibodies (Jackson Immoresearch). For quantitation of ovarian follicles and pituitary cells, sections were viewed using a Zeiss Axiocvert S100 microscope and video camera. Each section was covered systematically by accumulating non-overlapping fields of $1.5 \times 10^6 \mu\text{m}^2$. Counts and analysis were performed using Openlab image analysis software (Improvision Imaging).

Ovariectomy and sex steroid treatment

Ovariectomies were performed on 6-week-old virgins by veterinary services at Taconics Inc. Mice were allowed 2 weeks to recover from surgery. Subsequently, animals were given daily subcutaneous injections (100 µl) of 1 mg oestradiol or 1 µg progesterone, or a combination of both hormones for a period of 3 weeks. An immunolysis of these hormones for injection was prepared in mineral oil. Control animals were injected daily with vehicle.

Superovulation and oocyte retrieval

We induced 6-week-old virgins to ovulate by hormonal treatment. Animals were injected intraperitoneally with 10 units of pregnant mare horse serum (Boehringer-Mannheim) in sterile saline, and 35–40 h later with 15 units of human gonadotrophic hormone (Sigma). We collected oocytes from oviducts the following morning.

In vivo hormone stimulations and immunoblotting

For *in vivo* stimulations with either insulin or leptin, animals were fasted overnight (14–16 h). After anesthesia with sodium amobarbital (100 mg kg⁻¹, intraperitoneal), a bolus of insulin (5 units regular human insulin) or leptin (0.5 µg g⁻¹ body weight) was injected through the inferior vena cava. Controls received a comparable amount of diluent. Tissues (ovary, uterus and pituitary) were removed at times consistent with hormone action and homogenized at 4 °C as described⁸. Homogenates were allowed to solubilize for 1 h on ice and clarified by centrifugation at 16,000 g for 30 min. For detection of insulin-stimulated tyrosine phosphorylation, supernatants containing total protein (2 mg) were immunoprecipitated with an anti-IRS-1 or anti-IRS-2 antibodies⁸. Immunoblots were probed with anti-phosphotyrosine antibodies and detected by ¹²⁵I-protein A (ICN biochemicals). Blots were subsequently stripped and re-probed to reveal expression of IRS proteins. To assess activation of leptin pathways, we incubated hypothalamic lysates (0.5 mg) with antibodies to STAT3 (New England Biolabs) and then probed immunocomplexes with phospho-specific antibodies (New England Biolabs). For detection of the oestrogen receptor, 100 µg of tissue lysates was separated by SDS–PAGE and transferred to nitrocellulose. Membranes were probed with mouse monoclonal anti-ER antibodies (Neomarkers). Detection was by ¹²⁵I-protein A.

Metabolic measurements

Blood glucose levels were determined from mouse tails using a Glucometer Elite (Bayer). Glucose tolerance tests were carried out on mice after a 16-h fast as described⁸.

Food intake and body composition analysis

Body weight was measured twice each week beginning on postnatal day 21. For measurement of food consumption, 6-week-old female mice were housed individually and food intake and body weight were assessed daily between 9:00 and 11:00 for a period of 14 days. Water and food were available *ad libitum*. Food consumption was deduced from the weight of chow remaining in wire-mesh feeding jars. Animals received a standard laboratory diet (Purina Mouse Chow) consisting of 6.5% (wt/wt) fat, 47% carbohydrates, and 23.5% protein. Animals were killed at 8 weeks of age and blood was obtained for metabolic measurements (leptin, insulin, triglycerides) by cardiac puncture. Carcasses were digested by alcoholic potassium hydroxide hydrolysis at 60 °C overnight and body lipid was analysed as described³⁰.

Received 11 April; accepted 14 June 2000.

- Franks, S., Gilling-Smith, C., Watson, H. & Willis, D. Insulin action in the normal and polycystic ovary. *Endocrin. Metab. Clin. North Am.* **28**, 361–378 (1999).
- Legro, R. S. *et al.* Phenotype and genotype in polycystic ovary syndrome. *Recent Prog. Horm. Res.* **53**, 217–256 (1998).
- Tissenbaum, H. A. & Ruvkun, G. An insulin-like signaling pathway affects both longevity and reproduction in *Caenorhabditis elegans*. *Genetics* **148**, 703–717 (1998).
- Hsin, H. & Kenyon, C. Signals from the reproductive system regulate the lifespan of *C. elegans*. *Nature* **399**, 362–366 (1999).
- Bohni, R. *et al.* Autonomous control of cell and organ size by CHICO, a *Drosophila* homolog of vertebrate IRS1–4. *Cell* **97**, 865–875 (1999).
- Lavan, B. E., Lane, W. S. & Lienhard, G. E. The 60-kDa phosphotyrosine protein in insulin-treated adipocytes is a new member of the insulin receptor substrate family. *J. Biol. Chem.* **272**, 11439–11443 (1997).
- Lavan, B. E. *et al.* A novel 160 kDa phosphotyrosine protein in insulin-treated embryonic kidney cells is a new member of the insulin receptor substrate family. *J. Biol. Chem.* **272**, 21403–21407 (1997).
- Withers, D. J. *et al.* Disruption of IRS-2 causes type 2 diabetes in mice. *Nature* **391**, 900–904 (1998).
- Withers, D. J. *et al.* Irs-2 coordinates Igf-1 receptor-mediated beta-cell development and peripheral insulin signalling. *Nature Genet.* **23**, 32–40 (1999).
- Ogg, S. *et al.* The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. *Nature* **389**, 994–999 (1997).
- Morita, Y. *et al.* Requirement for phosphatidylinositol-3-kinase in cytokine-mediated germ CELL survival during fetal oogenesis in the mouse. *Endocrinology* **140**, 941–949 (1999).
- Vassen, L., Wegrzyn, W. & Klein-Hitpass, L. Human insulin receptor substrate-2 (IRS-2) is a primary progesterone response gene. *Mol. Endocrinol.* **13**, 485–494 (1999).
- Richards, R. G., DiAugustine, R. P., Petrusz, P., Clark, G. C. & Sebastian, J. Estradiol stimulates tyrosine phosphorylation of the insulin-like growth factor-1 receptor and insulin receptor substrate-1 in the uterus. *Proc. Natl Acad. Sci. USA* **93**, 12002–12007 (1996).
- Schwartz, M. W., Baskin, D. G., Kaiyala, K. J. & Woods, S. C. Model of the regulation of energy balance and adiposity by the central nervous system. *Am. J. Clin. Nutr.* **69**, 584–596 (1999).
- Unger, J. W. & Betz, M. Insulin receptors and signal transduction proteins in the hypothalamo–hypophyseal system: a review on morphological findings and functional implications. *Histol. Histopathol.* **13**, 1215–1224 (1998).
- Unger, J. W. & Lange, W. Insulin receptors in the pituitary gland: morphological evidence for influence on opioid peptide-synthesizing cells. *Cell Tissue Res.* **288**, 471–483 (1997).
- Woods, S. C., Stein, L. J., McKay, L. D. & Porte, D. Chronic intracerebroventricular infusion of insulin reduces food intake and body weight of baboons. *Nature* **282**, 503–505 (1979).
- Foster, L. A., Ames, N. K. & Emery, R. S. Food intake and serum insulin responses to intraventricular infusions of insulin and IGF-1. *Physiol. Behav.* **50**, 745–749 (1991).
- Flier, J. S. What's in a name? In search of leptin's physiological role. *J. Clin. Endocrin. Metab.* **83**, 1407–1413 (1998).
- Schwartz, M. & Seeley, R. Neuroendocrine responses to starvation and weight loss. *New Engl. J. Med.* **336**, 1802–18011 (1997).
- Ahima, R., Prabakaran, D., Mantzoros, C. *et al.* Role of leptin in the neuroendocrine response to fasting. *Nature* **382**, 250–252 (1996).
- Chehab, F. F., Lim, M. E. & Lu, R. Correction of the sterility defect in homozygous obese female mice by treatment with the human recombinant leptin. *Nature Genet.* **12**, 318–320 (1996).
- Chen, H. *et al.* Evidence that the diabetes gene encodes the leptin receptor: Identification of a mutation in the leptin receptor gene in db/db mice. *Cell* **84**, 491–495 (1996).
- Lee, G. H. *et al.* Abnormal splicing of the leptin receptor in *diabetic* mice. *Nature* **379**, 632–635 (1996).
- Vaisse, C. *et al.* Leptin activation of Stat3 in the hypothalamus of wild-type and *ob/ob* mice but not *db/db* mice. *Nature Genet.* **14**, 95–97 (1996).
- McCowen, K. C., Chow, J. C. & Smith, R. J. Leptin signaling in the hypothalamus of normal rats *in vivo*. *Endocrinology* **139**, 4442–4447 (1998).
- Szanto, I. & Kahn, C. R. Selective interaction between leptin and insulin signaling pathways in a hepatic cell line. *Proc. Natl Acad. Sci.* **97**, 2355–2360 (2000).
- Berti, L., Kellerer, M., Capp, E. & Haring, H. U. Leptin stimulates glucose transport and glycogen synthesis in C2C12 myotubes: evidence for a PI 3-kinase mediated effect. *Diabetologia* **40**, 606–609 (1997).
- Cohen, B., Novick, D. & Rubinstein, M. Modulation of insulin activities by leptin. *Science* **274**, 1185–1188 (1996).
- Fredrich, R. C. *et al.* Leptin levels reflect body lipid content in mice: evidence for diet-induced resistance to leptin action. *Nature Med.* **1**, 1311–1314 (1995).

Acknowledgements

We are grateful to C. Cahill for preparation of pituitary sections; A. Parlow for antibodies to pituitary hormones; A. Dunaif for testosterone measurements; and D. Hess for helpful discussions of murine sex hormones.

Correspondence and requests for materials should be addressed to M.F.W. (e-mail: morris.white@joslin.harvard.edu).



Cell-fate conversion of lymphoid-committed progenitors by instructive actions of cytokines

Motonari Kondo*, David C. Scherer*, Toshihiro Miyamoto*, Angela G. King*, Koichi Akashi*, Kazuo Sugamura† & Irving L. Weissman*

* Departments of Pathology and Developmental Biology, Stanford University School of Medicine, Stanford, California 94305-5324, USA

† Department of Microbiology and Immunology, Tohoku University School of Medicine, Sendai 980-8575, Japan

The primary role of cytokines in haemato-lymphopoiesis is thought to be the regulation of cell growth and survival¹⁻³. But the instructive action of cytokines in haematopoiesis has not been well addressed⁴. Here we show that a clonogenic common lymphoid progenitor⁵, a bone marrow-resident cell that gives rise exclusively to lymphocytes (T, B and natural killer cells), can be redirected to the myeloid lineage by stimulation through exogenously expressed interleukin (IL)-2 and GM-CSF (granulocyte/macrophage colony-stimulating factor) receptors. Analysis of mutants of the β -chain of the IL-2 receptor revealed that the granulocyte- and monocyte-differentiation signals are triggered by different cytoplasmic domains, showing that the signalling pathway(s) responsible for these unique developmental outcomes are separable. Finally, we show that the endogenous myelo-monocytic cytokine receptors for GM-CSF and macrophage colony-stimulating factor (M-CSF) are expressed at low to moderate levels on the more primitive haematopoietic stem cells, are absent on common lymphoid progenitors, and are upregulated after myeloid lineage induction by IL-2. We conclude that cytokine signalling can regulate cell-fate decisions and propose that a critical step in lymphoid commitment is downregulation of

cytokine receptors that drive myeloid cell development.

IL-15 has been proposed to be a critical cytokine for natural killer (NK) cell development⁶⁻⁸. The receptor for IL-15 shares two subunits with the IL-2 receptor (IL-2R), namely the common γ (γ_c) and IL-2R β chains^{9,10}. Given this conservation of subunits, we decided to determine whether IL-2 receptor signalling can induce NK cell development from common lymphoid progenitors (CLPs). CLPs express the γ_c but not the IL-2R β chain (data not shown). Human IL-2R β can couple with mouse γ_c chain to form a functional IL-2 receptor that responds exclusively to human IL-2 (hIL-2)^{11,12}; therefore, we cultured CLPs from mice transgenic for the human IL-2R β chain¹³ on OP9 bone marrow stromal cells¹⁴ in the presence of hIL-2. IL-2R β transgene expression is driven by the MHC class I promoter and cell-surface expression was observed in nearly all haematopoietic cells, including haematopoietic stem cells (HSCs) and CLPs (data not shown). After 10 days the culture contained an unexpectedly large number of Gr-1⁺ granulocytes and monocytes (GM lineage cells, Fig. 1a) in addition to CD19⁺ B cells and NK1.1⁺ NK cells (see Fig. 2). Under these same culture conditions, CLPs from wild-type mice (C57BL/Ka-Thy1.1) gave rise to B and NK cells but never to GM cells (data not shown). In the absence of hIL-2, CLPs from IL-2R β transgenic mice are still lymphoid-committed, even in the presence of the myelo/monocytic cytokines IL-3, GM-CSF and M-CSF (Fig. 1a). When we carried out colony-forming assays in methylcellulose, we found that in addition to CLPs, pro-T cells from the thymus of IL-2R β transgenic mice could also convert to the myeloid lineage when stimulated with hIL-2 (Fig. 1b). However, IL-2R β -expressing pro-B cells from the bone marrow did not give rise to GM cells under any culture conditions. These data indicate that CLPs and pro-T cells may have a latent GM lineage differentiation program that can be initiated by signalling through the reconstituted IL-2 receptor. It is important to note that there is a 2-day window after which cultured CLPs lose this IL-2-induced GM lineage differentiation potential and 'irreversibly' commit to the lymphoid lineage (Fig. 1a). Conversely, the lineage converting effect of IL-2 is complete in 2 days and can be withdrawn without any deleterious effect on the conversion process under GM growth

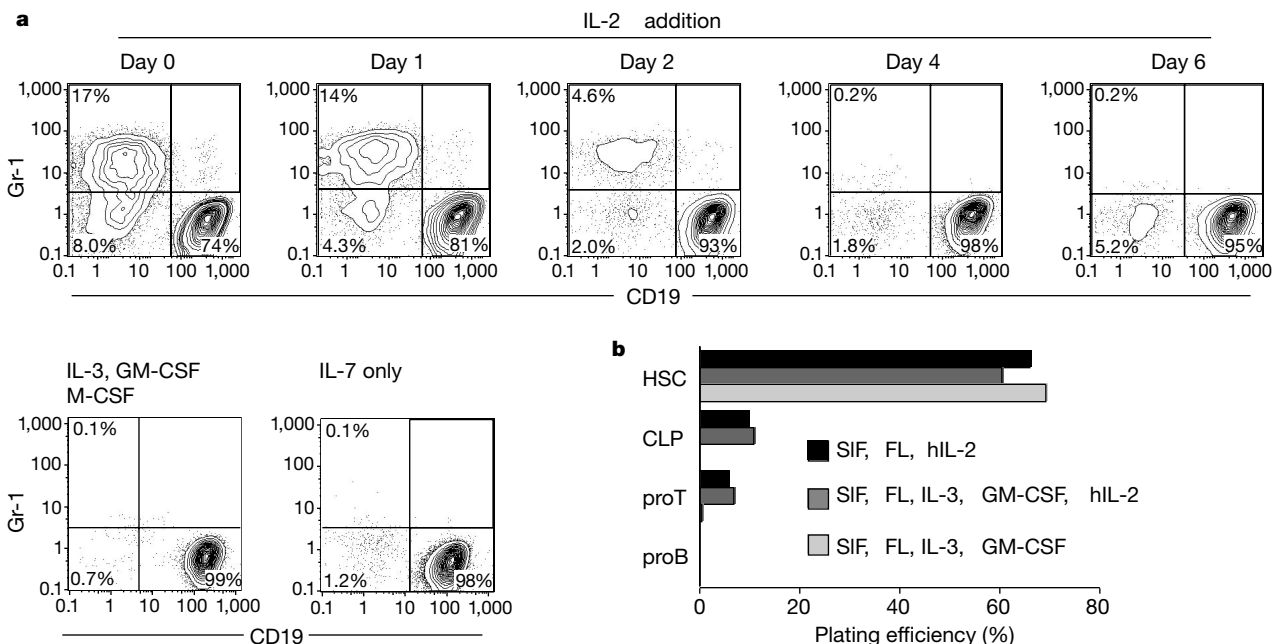


Figure 1 Granulocyte/macrophage differentiation from CLPs of IL-2R β transgenic mice. **a**, Flow cytometric analysis of cell populations from 40 CLPs of IL-2R β transgenic mice after stromal cell culture with OP9 in the presence of IL-7 (10 ng ml⁻¹) for 10 days. Human IL-2 (25 ng ml⁻¹) was added to the culture at the indicated time points (top row). No Gr-1⁺ cells were observed when GM-CSF (10 ng ml⁻¹), IL-3 (30 ng ml⁻¹) and M-CSF (25 U ml⁻¹)

were added into the culture (bottom row, left). **b**, Colony-forming activity of cells from IL-2R β transgenic mice. FACS-sorted cells were cultured in methylcellulose for 5–8 days in the presence of various cytokines indicated in the figure. All colonies were morphologically GM colonies. Results shown are means of more than three independent cultures.

conditions (data not shown). In addition, we found that the cell-fate conversion was limited to the GM lineage, as no colonies grew under erythroid culture conditions with IL-2.

There are two possible explanations for the GM lineage readout that we observed in our experimental system. First, IL-2 receptor signalling may be rescuing from apoptosis a myeloid-committed population of cells that co-purifies with CLPs. However, we observed no lineage conversion from CLPs expressing the anti-apoptotic protein Bcl-2 (ref. 15) when cultured under GM growth conditions (data not shown). Second, IL-2 receptor signalling may be actively promoting the development of GM lineage cells from lymphoid-committed progenitors. To test this hypothesis, we examined the trans-differentiation of CLPs from IL-2R β transgenic mice at the clonal level.

Limiting dilution analysis¹⁶ revealed that the GM cell readout frequency of CLPs from IL-2R β transgenic mice was 1 event per 4 cells when IL-2 was added at day 0 (Table 1). In the absence of IL-2, B-cell readout frequency of CLPs was 1 out of 10. However, B-cell development dropped to the level of 1 in 40 when IL-2 was added at

day 0, suggesting that GM lineage conversion had occurred at the expense of lymphoid development.

Next, we clone-sorted CLPs from IL-2R β transgenic mice into 96-well plates and added IL-2 at day 0, 2 or 4 of the culture. Consistent with the results of the limiting dilution assay above, B-cell readout was severely blocked when IL-2 was added into the culture at day 0 or 2, when GM lineage conversion can be initiated (data not shown; see Supplementary Information). More importantly, about 20% of scorable wells had both GM and lymphoid lineage cells when IL-2 was added at day 2 (Fig. 2; see GM/B, GM/NK and GM/B/NK). This directly shows that IL-2 receptor signalling can drive myeloid lineage commitment from lymphoid committed progenitors. In addition to this observation, ~5% of pro-T-cell-derived GM colonies from IL-2R β transgenic mice (Fig. 1b) had rearranged T-cell receptor loci, an event seen only in lymphoid lineage cells¹⁷ (A.G.K., M.K. and I.L.W., manuscript in preparation). These results show that early lymphoid committed progenitors maintain a latent GM lineage differentiation activity that can be triggered by signalling through an exogenously expressed cytokine receptor.

To assess the molecular mechanisms that drive GM lineage conversion in lymphoid-restricted progenitors, we used retrovirus-mediated gene transfer to introduce wild-type and mutant human IL-2R β chains¹⁸ into CLPs from wild-type mice and

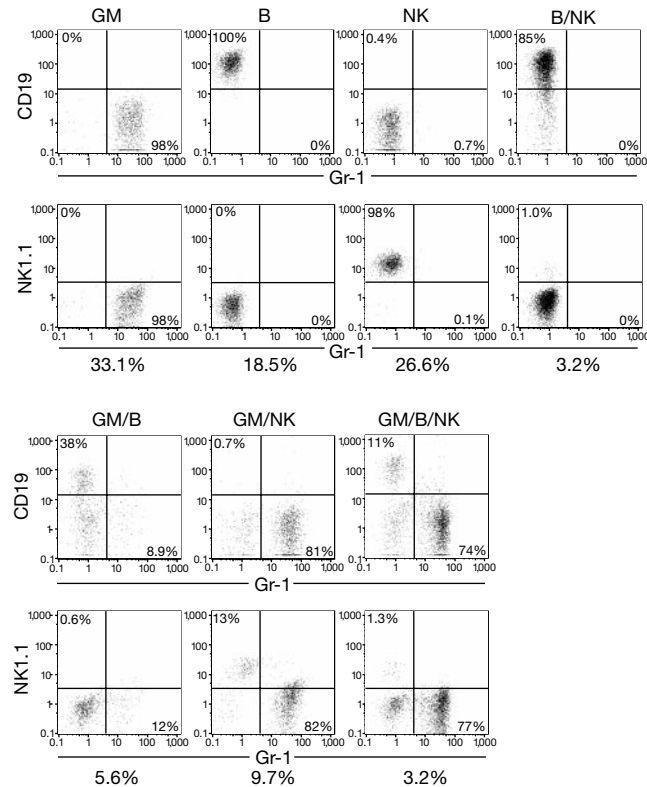


Figure 2 Clonogenic trans-differentiation activity of CLPs from IL-2R β transgenic mice. Single ACDU-sorted CLPs from IL-2R β transgenic mice were each cultured on an OP9 stromal cell layer in each well of 96-well plates for 7 days in the presence of SIF, FL and IL-7. Human IL-2 was added at day 2 of culture. Positive readout wells were determined by microscopic observation (124 wells out of 576 wells, 21.5%). The readout pattern was determined by flow cytometric analysis. The frequency of wells with each differentiation pattern is shown as a percentage of total positive wells under each FACS plot.

Table 1 Limiting number of IL-2R β transgenic CLPs for each lineage readout

Readout cells	Day of culture when human IL-2 was added		
	Day 0	Day 4	—
GM cells	4	>500	∞
B cells	40	11	10
NK cells	27	24	150

Between 1 and 40 CLPs from IL-2R β transgenic mice were cultured on an OP9 stromal cell layer in the presence of IL-7. Human IL-2 was added at indicated time points. The graphs used to obtain the limiting numbers are given in the Supplementary Information.

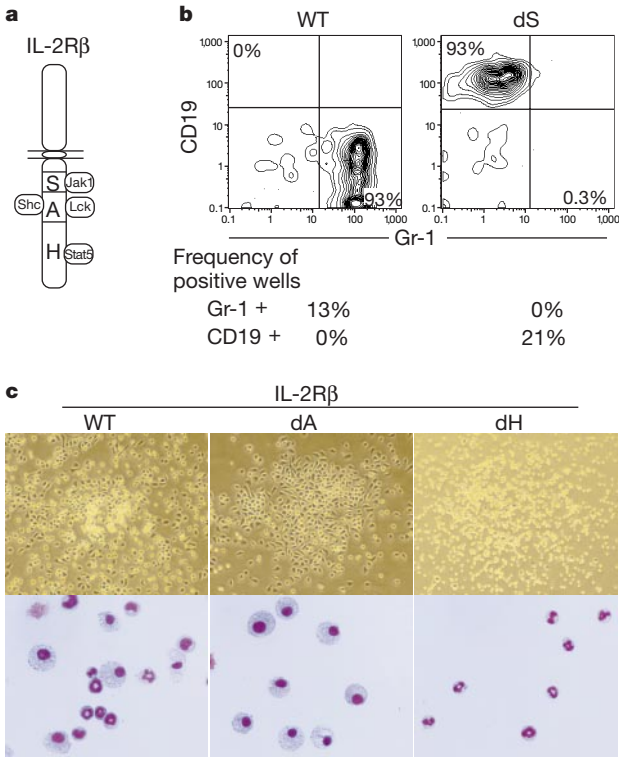


Figure 3 Induction of myeloid cells from CLPs of wild-type mice. **a**, Representation of IL-2R β structure. See Supplementary Information for diagram of mutants. **b**, CLPs from C57BL/Ka-Thy1.1 mice were infected with either wild-type IL-2R β (WT) or the dS mutant retrovirus. After culture in the presence of IL-2 for 48 h, GFP⁺ cells were sorted and cultured on OP9 cells for 6 days in the presence of IL-3 and GM-CSF. Gr-1⁺ cells were morphologically granulocytes and monocytes. **c**, Lineage conversion induced through IL-2R β mutants. Wild-type, dA and dH mutants of IL-2R β chains were introduced in CLPs from wild-type mice by retroviral infection. After pre-culture as in **b**, GFP⁺ cells were further cultured in methylcellulose for 5–8 days in the presence of SIF, IL-3 and GM-CSF (see Supplementary Information). Colony type was determined under microscopic observation and confirmed by cytopsin. Over 95% of cells in the colonies derived from CLPs with the dA and dH mutants were macrophages and granulocytes, respectively, by differential cell count.

analysed their development *in vitro*. All retroviral constructs contained a green fluorescent protein (GFP) complementary DNA driven by an internal ribosome entry site (IRES) for the purpose of identifying infected cells. After infection, we cultured cells for 48 h in the presence of hIL-2 to induce GM lineage conversion. We then sorted GFP⁺ CLPs, and cultured them on stromal cells for 6 days in the absence of hIL-2. CLPs expressing the wild-type IL-2R β chain gave rise to Gr-1⁺ cells but not to CD19⁺ B cells (Fig. 3b). In contrast, CLPs expressing the non-functional dS mutant¹⁸, which lacks the S region of the cytoplasmic tail of IL-2R β (Fig. 3a), differentiated into B cells, but not into GM lineage cells. This shows that signals emanating from the cytoplasmic tail of IL-2R β are critical for GM lineage conversion of CLPs.

To characterize further the signalling requirements for GM lineage conversion in CLPs, we analysed two additional IL-2R β mutants, dA and dH¹⁸, which lack the A and H regions of the cytoplasmic domain of IL-2R β chain, respectively (Fig. 3a). The A region is necessary for recruitment of Shc and Src tyrosine kinases, such as Lck and Fyn; the H region is crucial for Stat5 activation¹⁹. Infection and pre-culture of CLPs in the presence of hIL-2 was performed as before, and GFP⁺ cells were then sorted into methylcellulose culture medium that supports myeloid differentiation. After 5–7 days of culture, CLPs expressing wild-type IL-2R β formed GM colonies at 5–10% plating efficiency (Fig. 3c), whereas CLPs expressing the non-functional dS mutant did not form detectable colonies (data not shown). Notably, CLPs expressing the dA mutant formed exclusively macrophage colonies, whereas those with the dH mutant formed only granulocyte colonies (Fig. 3c). Thus

macrophage and granulocyte differentiation are induced by different cytoplasmic regions of the IL-2R β chain, and the signal transduction cascades that drive granulocyte and macrophage differentiation are separable.

The known signalling pathways activated by the IL-2 receptor have significant overlap with those activated by the receptors for erythropoietin (Epo) and IL-7. For example, Shc is phosphorylated in response to stimulation with either Epo or IL-2, whereas all three cytokines activate the transcription factor Stat5 (refs 19, 20). However, Epo stimulation did not induce GM lineage conversion in CLPs expressing retrovirally encoded EpoR (Fig. 4a). Likewise, IL-7 did not induce GM lineage conversion in CLPs through either endogenously or exogenously expressed IL-7 receptor. These data, together with the evidence obtained with the dA and dH mutants of IL-2R β , suggest that unknown signal transduction pathways are involved in the initiation of GM lineage conversion in CLPs. They also show that GM lineage conversion in CLPs is the result of the activation of specific signalling cascades for differentiation, not just the supportive effect of cytokines for cell proliferation and survival.

What is the significance of IL-2 induced GM lineage conversion of CLPs expressing IL-2R β ? One possibility is that the IL-2 receptor mimics other cytokine receptors that support GM lineage differentiation, such as the GM-CSF receptor. To examine this possibility, we transduced wild-type CLPs with a retrovirus that expresses a functional human GM-CSF receptor²¹ and cultured these cells in the presence of human GM-CSF for 2 days. After infection, the cells were then replated in methylcellulose under myeloid culture conditions. After 5–7 days in culture, GM colonies were formed (Fig. 4b) in transduced but not in wild-type CLPs. Expression of the exogenous human GM-CSF receptor in the colony-forming cells was confirmed by polymerase chain reaction with reverse transcription (RT–PCR) (Fig. 4b). This result prompted us to examine the expression of the GM-CSF and M-CSF receptors in CLPs from IL-2R β transgenic mice in response to IL-2. In the absence of IL-2, CLPs expressed neither the GM-CSF receptor α -chain nor the M-CSF receptor (Fig. 4c); however, expression of these two molecules was strongly induced in the presence of IL-2 after 48 h. HSCs, which give rise to both lymphoid and myeloid lineages, expressed low to moderate levels of the GM-CSF and M-CSF receptors. From these data, we propose that one of the earliest events in lymphoid lineage commitment is downregulation of cytokine receptors whose function is to drive and support GM lineage differentiation.

Significantly, some acute myeloid leukemia (AML) cells have rearrangements of immunoglobulin and T-cell receptor genes together with myeloid-specific antigens^{22–24}. The developmental plasticity of lymphoid lineage progenitors revealed here may help to explain such 'lineage infidelities'. Finally, the finding that cytokine receptor signalling can instructively drive lineage commitment will allow us to begin to elucidate the molecular mechanisms that regulate haematopoietic cell differentiation. □

Methods

Mice

Wild-type mice are C57BL/Ka-Thy1.1. Human IL-2R β transgenic mice, originally bred on the C57BL/6 background¹³, were backcrossed to C57BL/Ka-Thy1.1 over two generations. Mice were bred and maintained in the animal care facility at Stanford University School of Medicine and were used at 5–10 weeks of age.

Flow cytometric analysis and cell sorting

We purified CLPs as described⁵. The surface phenotypes of thymic pro-T and bone marrow pro-B cells used in this study are CD3⁺CD8⁺c-Kit^{high}IL-2R α IL-7R α ⁺ (ref. 25) and IgM⁺B220⁺CD43⁺HSA⁺NK1.1⁺ (refs 26, 27), respectively. All cell sorting and flow cytometric analysis were done by a highly modified double or triple laser FACS Vantage available at the Stanford shared FACS facility.

Plasmid construction

We cloned cDNAs of human IL-2R β and its mutants into the EcoRI site of the retroviral vector, MSCV-IRES–GFP (MSCV-2R β -IRES–GFP, MSCV-dS-IRES–GFP, MSCV-dA-

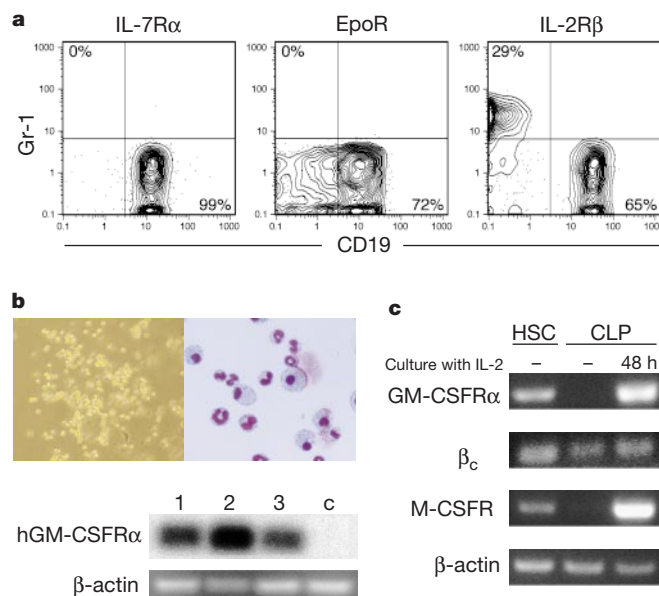


Figure 4 Granulocyte/macrophage induction in CLPs is initiated by stimulation with GM-CSF. **a**, Epo stimulation does not induce GM lineage conversion in CLPs. IL-7R α (left), EpoR (middle) and IL-2R β (right) were introduced into CLPs from C57BL/Ka-Thy1.1 mice by retroviral vectors. For infection and pre-culture, see Methods and Supplementary Information. GFP⁺ cells were sorted and cultured on OP9 stromal cells in the presence of IL-7 (left), Epo (middle) and IL-2 (right), as well as SIF and FL. After 7 days in the stromal cell culture, readout cell populations were examined by a flow cytometer. NK1.1⁺ cells were observed in some cases. The plots shown are results after gating on the NK1.1⁺ population. **b**, GM colony formation from CLPs by stimulation with GM-CSF. CLPs from wild-type mice were transduced with human GM-CSF receptor by a retroviral vector. Cell culture was done as described in Methods. After 5–8 days in methylcellulose culture, GM colonies were observed. Retrovirally introduced gene expression in three independent GM colonies (lanes 1–3) was confirmed by RT–PCR. **c**, Expression of the indicated genes was determined by RT–PCR in HSCs and CLPs from IL-2R β transgenic mice in the presence or absence of human IL-2 (48-h stimulation).

IRE5–GFP and MSCV-dH-IRE5–GFP). The dS, dA and dH mutants lack the S region (amino acid residues 267–322), the A region (residues 313–382) and H region (residues 383–525), respectively. We created MSCV-IL-7R α -IRE5–GFP by inserting a mouse IL-7R α cDNA into the *Xho*I site of MSCV-IRE5–GFP. To generate the MSCV-GMR α -IRE5– β _c vector, we inserted human GMR α cDNA into the *Xho*I site of the MSCV-IRE5–GFP vector (MSCV-GMR α -IRE5–GFP) and put the human β _c cDNA downstream of the IRES in place of the GFP cDNA. A retroviral vector for EpoR (MSCV-EpoR-IRE5–GFP)²⁸ was provided by G. Q. Daley.

Retroviral gene transfer

We produced virus from a packaging cell line, Phoenix, by transient transfection of MSCV constructs with FuGENE 6 (Roche Molecular Biochemicals). For viral infection, CLPs were cultured at 5,000 cells per well in a 96-well round-bottomed plate with 50% virus stock in Iscove's Modified Dulbecco's Media (IMDM) containing 10% FCS, 50 μ M 2-mercaptoethanol, 4 μ g ml⁻¹ polybrene, SIF (20 ng ml⁻¹), FL (20 ng ml⁻¹) and IL-7 (10 ng ml⁻¹) for 12 h at 32 °C. Cells were further cultured for 36–48 h at 37 °C to allow exogenous gene expression. Human IL-2 (25 ng ml⁻¹), erythropoietin (4 U ml⁻¹) and GM-CSF (20 ng ml⁻¹) were used in this culture for stimulating the human IL-2R β , EpoR- and GM-CSFR-expressing CLPs, respectively.

In vitro cell culture

Between 1 and 40 CLPs were sorted by the automatic cell deposition unit (ACDU, Beckton Dickinson) on the FACS Vantage. The cells were cultured in 96-well plates on an OP9 stromal layer in the presence of indicated cytokines for 7–10 days. Methylcellulose culture was done as described⁵.

RT–PCR analysis

We isolated RNA from cells with Trizol reagent (GIBCO BRL). Oligo-dT primed cDNA was subjected to PCR. After electrophoresis with 2% agarose gel, bands were detected by ethidium bromide staining. Amplified bands for human GM-CSFR α were detected by Southern blotting with the radiolabelled synthesized oligo as a probe. For the primer sequence, see Supplementary Information.

Receive 22 May; accepted 31 July 2000.

1. Fairbairn, L. J., Cowling, G. J., Reipert, B. M. & Dexter, T. M. Suppression of apoptosis allows differentiation and development of a multipotent hemopoietic cell line in the absence of added growth factors. *Cell* **74**, 823–832 (1993).
2. Ogawa, M. Differentiation and proliferation of hematopoietic stem cells. *Blood* **81**, 2844–2853 (1993).
3. Watowich, S. S. *et al.* Cytokine receptor signal transduction and the control of hematopoietic cell development. *Annu. Rev. Cell Dev. Biol.* **12**, 91–128 (1996).
4. Metcalf, D. Stem cells, pre-progenitor cells and lineage-committed cells: are our dogmas correct? *Ann. NY Acad. Sci.* **872**, 289–303 (1999); discussion **872**, 303–304 (1999).
5. Kondo, M., Weissman, I. L. & Akashi, K. Identification of clonogenic common lymphoid progenitors in mouse bone marrow. *Cell* **91**, 661–672 (1997).
6. Ogasawara, K. *et al.* Requirement for IRF-1 in the microenvironment supporting development of natural killer cells. *Nature* **391**, 700–703 (1998); erratum **392**, 843 (1998).
7. Lodolce, J. P. *et al.* IL-15 receptor maintains lymphoid homeostasis by supporting lymphocyte homing and proliferation. *Immunity* **9**, 669–676 (1998).
8. Kennedy, M. K. *et al.* Reversible defects in natural killer and memory CD8 T cell lineages in interleukin 15-deficient mice. *J. Exp. Med.* **191**, 771–780 (2000).
9. Giri, J. G. *et al.* Utilization of the beta and gamma chains of the IL-2 receptor by the novel cytokine IL-15. *EMBO J.* **13**, 2822–2830 (1994).
10. Sugamura, K. *et al.* The interleukin-2 receptor gamma chain: its role in the multiple cytokine receptor complexes and T cell development in XSCID. *Annu. Rev. Immunol.* **14**, 179–205 (1996).
11. Kondo, M. *et al.* Sharing of the interleukin-2 (IL-2) receptor gamma chain between receptors for IL-2 and IL-4. *Science* **262**, 1874–1877 (1993).
12. Nemoto, T. *et al.* Differences in the interleukin-2 (IL-2) receptor system in human and mouse: alpha chain is required for formation of the functional mouse IL-2 receptor. *Eur. J. Immunol.* **25**, 3001–3005 (1995).
13. Asano, M. *et al.* IL-2 can support growth of CD8+ T cells but not CD4+ T cells of human IL-2 receptor beta-chain transgenic mice. *J. Immunol.* **153**, 5373–5381 (1994).
14. Nakano, T., Kodama, H. & Honjo, T. Generation of lymphohematopoietic cells from embryonic stem cells in culture. *Science* **265**, 1098–1101 (1994).
15. Domen, J., Gandy, K. L. & Weissman, I. L. Systemic overexpression of BCL-2 in the hematopoietic system protects transgenic mice from the consequences of lethal irradiation. *Blood* **91**, 2272–2282 (1998).
16. Henry, C., Marbrook, J., Vann, D. C., Kodlin, D. & Wofsy, C. In *Selected Methods in Cellular Immunology* (eds Mishell, B. B. & Shiigi, S. M.) 138–152 (WH Freeman, San Francisco, 1981).
17. Schlissel, M. S. & Stanhope-Baker, P. Accessibility and the developmental regulation of V(D)J recombination. *Semin. Immunol.* **9**, 161–170 (1997).
18. Hatakeyama, M., Mori, H., Doi, T. & Taniguchi, T. A restricted cytoplasmic region of IL-2 receptor beta chain is essential for growth signal transduction but not for ligand binding and internalization. *Cell* **59**, 837–845 (1989).
19. Nelson, B. H. & Willerford, D. M. Biology of the interleukin-2 receptor. *Adv. Immunol.* **70**, 1–81 (1998).
20. Constantinescu, S. N., Ghaffari, S. & Lodish, H. F. The erythropoietin receptor: structure, activation and intracellular signal transduction. *Trends Endocrinol. Metab.* **10**, 18–23 (1999).
21. Watanabe, S. *et al.* Reconstituted human granulocyte-macrophage colony-stimulating factor receptor transduces growth-promoting signals in mouse NIH 3T3 cells: comparison with signalling in BA/F3 pro-B cells. *Mol. Cell. Biol.* **13**, 1440–1448 (1993).
22. Ha, K., Minden, M., Hozumi, N. & Gelfand, E. W. Immunoglobulin gene rearrangement in acute myelogenous leukemia. *Cancer Res.* **44**, 4658–4660 (1984).

23. Palumbo, A., Minowada, J., Erikson, J., Croce, C. M. & Rovera, G. Lineage infidelity of a human myelogenous leukemia cell line. *Blood* **64**, 1059–1063 (1984).
24. Cheng, G. Y., Minden, M. D., Toyonaga, B., Mak, T. W. & McCulloch, E. A. T cell receptor and immunoglobulin gene rearrangements in acute myeloblastic leukemia. *J. Exp. Med.* **163**, 414–424 (1986).
25. Godfrey, D. I., Kennedy, J., Mombaerts, P., Tonegawa, S. & Zlotnik, A. Onset of TCR-beta gene rearrangement and role of TCR-beta expression during CD3⁺ CD4⁺ CD8⁺ thymocyte differentiation. *J. Immunol.* **152**, 4783–4792 (1994).
26. Hardy, R. R., Carmack, C. E., Shinton, S. A., Kemp, J. D. & Hayakawa, K. Resolution and characterization of pro-B and pre-pro-B cell stages in normal mouse bone marrow. *J. Exp. Med.* **173**, 1213–1225 (1991).
27. Rolink, A. *et al.* A subpopulation of B220+ cells in murine bone marrow does not express CD19 and contains natural killer cell progenitors. *J. Exp. Med.* **183**, 187–194 (1996).
28. Ghaffari, S. *et al.* BCR-ABL and v-SRC tyrosine kinase oncoproteins support normal erythroid development in erythropoietin receptor-deficient progenitor cells. *Proc. Natl Acad. Sci. USA* **96**, 13186–13190 (1999).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank T. Honjo for IL-2R β transgenic mice; G. Nolan for the retroviral system; S. Watanabe for cDNAs for human GM-CSFR α and β _c; G. Q. Daley for a retroviral expression vector for EpoR; S.-I. Nishikawa for OP9 cells and anti-IL-7R α antibody; and L. Jerabek for laboratory management. This work was supported by USPHS grant to I.L.W. and a Jose Carreras International Leukemia Foundation Grant to K.A. M.K., D.C.S. and A.G.K. are supported by fellowships from the Irvington Institute for Immunology, American Cancer Society California Division, and USPHS Training Grant, respectively.

Correspondence and requests for materials should be addressed to M.K. (e-mail: motonari.kondo@stanford.edu).

Tat-specific cytotoxic T lymphocytes select for SIV escape variants during resolution of primary viraemia

Todd M. Allen^{*†}, David H. O'Connor^{*†}, Peicheng Jing^{*}, John L. Dzuris[‡], Bianca R. Mothé^{*}, Thorsten U. Vogel^{*}, Ed Dunphy^{*}, Max E. Liebl^{*}, Carol Emerson^{*}, Nancy Wilson^{*}, Kevin J. Kunstman^{||}, Xiaochi Wang[¶], David B. Allison[‡], Austin L. Hughes[✱], Ronald C. Desrosiers^{**}, John D. Altman[¶], Steven M. Wolinsky^{||}, Alessandro Sette[‡] & David I. Watkins^{*††}

^{*} Wisconsin Regional Primate Research Center, University of Wisconsin, 1220 Capitol Court, Madison, Wisconsin 53715-1299, USA
[‡] Epimmune, 5820 Nancy Ridge Drive, San Diego, California 92121, USA
^{||} Northwestern University Medical School, 303 East Chicago, Room 3-735 Tarry Building, Chicago, Illinois 60611-3008, USA
[¶] Emory Vaccine Center, Emory University School of Medicine, G211 Rollins Research Building, 1510 Clifton Road, Atlanta, Georgia 30322, USA
[#] St. Luke's/Roosevelt Hospital, Obesity Research Center, 1090 Amsterdam Avenue, Suite 14B, New York, New York 10025, USA
[✱] Department of Biological Sciences, 401 Coker Life Sciences, University of South Carolina, Columbia, South Carolina 29208, USA
^{**} New England Regional Primate Research Center, One Pine Hill Drive, Southborough, Massachusetts 01772-9102, USA
^{††} Dept. of Pathology and Laboratory Medicine, University of Wisconsin, 1300 University Avenue, Madison, Wisconsin 53706-1532, USA
[†] These authors contributed equally to this work

Human immunodeficiency virus (HIV) and simian immunodeficiency virus (SIV) infections are characterized by early peaks of viraemia that decline as strong cellular immune responses develop^{1,2}. Although it has been shown that virus-specific CD8-positive cytotoxic T lymphocytes (CTLs) exert selective pressure during HIV and SIV infection^{3–11}, the data have been controversial^{12,13}. Here we show that Tat-specific CD8-positive T-lymphocyte responses select for new viral escape variants during

the acute phase of infection. We sequenced the entire virus immediately after the acute phase, and found that amino-acid replacements accumulated primarily in Tat CTL epitopes. This implies that Tat-specific CTLs may be significantly involved in controlling wild-type virus replication, and suggests that responses against viral proteins that are expressed early during the viral life cycle might be attractive targets for HIV vaccine development.

The strongest CD8-positive T lymphocyte responses to HIV and SIV are observed in the first few weeks of infection, coincident with the initial decline in plasma viraemia. We reasoned that viral escape might occur from immune responses that exert selective pressure during this acute phase of infection. To test this hypothesis, we examined viral evolution during acute infection of 18 rhesus macaques with molecularly cloned SIV.

Every animal (10/10) expressing the rhesus major histocompatibility (MHC) class I molecule Mamu-A*01 made CD8-positive T-lymphocyte responses, which peaked between 3 and 4 weeks after infection, to a newly defined epitope in Tat_{28–35} STPESANL¹⁴ (SL8, Fig. 1). In two of these animals, up to 10% of their CD8/CD3-positive lymphocytes recognized this Tat epitope. However, the frequency of Tat-specific lymphocytes declined precipitously after the acute phase (Fig. 1). We reasoned that this decline might be the result of viral escape from these Tat-specific responses. We investigated this possibility by sequencing the 5' exon of Tat using virus derived from the ten Mamu-A*01-positive animals. By 8 weeks after infection, a high frequency of amino-acid substitution was observed in the SL8 epitope (Fig. 2a). Eighty-six per cent of clones contained variation in the CD8-positive T-lymphocyte epitope at this time point. In five out of the ten Mamu-A*01-positive animals, all sequenced clones contained mutations in the Mamu-A*01-restricted SL8 epitope. In contrast, little amino-acid variation was observed outside this SL8 epitope in Mamu-A*01-positive animals (Fig. 2b).

We then investigated whether these changes in the SL8 epitope resulted from a mixed population of variants in our inocula or whether they were selected for increased replicative fitness in Mamu-A*01-negative animals. As expected from a molecular clone, there was little variation in this epitope in either of the two inocula (Fig. 2a). In addition, only one out of eight Mamu-A*01-

negative animals exhibited changes in the SL8 epitope (see Supplementary Information, Fig. 1). Thus, viral escape from the Mamu-A*01-restricted Tat-specific CD8-positive T-lymphocyte responses appeared to be the most consistent explanation for our findings.

We then performed a time-course analysis of viral evolution within the SL8 epitope and sequenced the entire virus after the acute phase in two of the Mamu-A*01-positive animals. At peak viraemia (2 weeks after infection), Tat-specific CTL responses were barely detectable and no changes in the Mamu-A*01-bound Tat epitope were present (Fig. 2c, d). After resolution of peak viraemia (3 weeks after infection), Tat-specific CD8-positive T lymphocytes were at their highest level. One week later, extensive variation was apparent in the virus populations of both animals (Fig. 2c, d). Furthermore, direct sequencing of the open reading frames of the entire virus at 4 weeks after infection revealed only a single site of viral nucleotide diversity in the SL8 epitope in animal 96118. In animal 96114 there were three sites of viral nucleotide diversity, one of which was in the SL8 epitope, and the other two in Rev and Env (see Supplementary Information; Fig. 2). In animal 96118, the nucleotide substitution in RNA encoding the SL8 epitope caused a change in the overlapping reading frame of Vpr. In animal 96114, the change in Rev also caused a substitution in the overlapping open reading frame of Env. This Rev replacement is seen in most animals infected with this viral clone and appears to be selected for increased viral fitness. Analysis of the additional replacement in animal 96114 in Env by interferon- γ enzyme-linked immunospot (ELISPOT) assays of CD8 and CD4 lymphocytes, however, failed to show conclusively that this region contained any T-cell epitopes.

To determine whether the observed sequence changes in the SL8 epitope do represent viral escape variants, we characterized the functional consequences of the predominant variant epitopes on peptide binding to Mamu-A*01 and on CTL recognition. *In vitro* peptide-binding analyses showed that the new variants of the SL8 epitope did not bind as well as the wild-type peptide to Mamu-A*01 (Table 1). The substitutions of proline at P1 and leucine at P5 reduced peptide binding by more than 50% and 80%, respectively. The isoleucine substitution at P2 and the glutamine, arginine and proline substitutions at P8 abrogated binding (> 99% reduction). As P2 is a secondary anchor and P8 is the carboxy anchor¹⁵ for peptides bound by the Mamu-A*01 molecule, we would expect

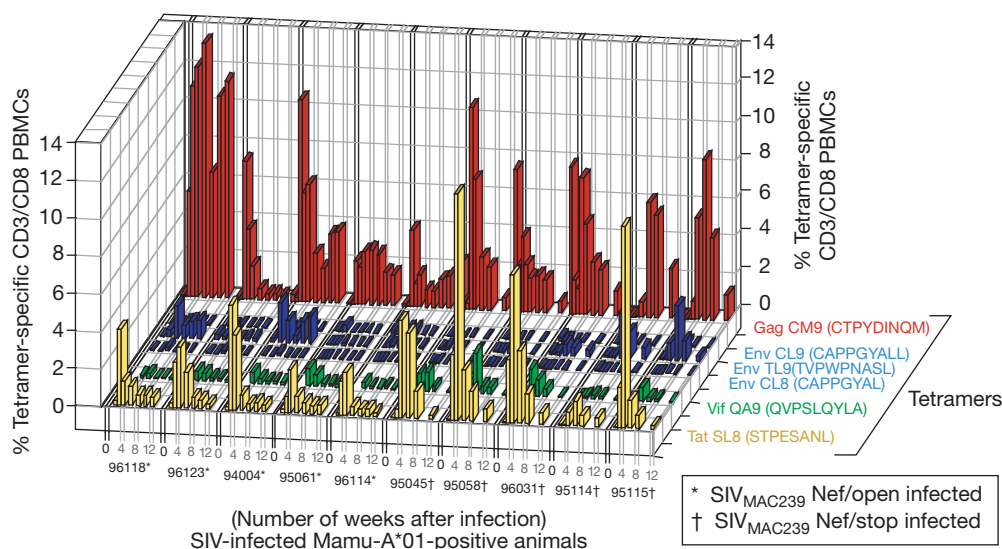


Figure 1 Quantitation of CD8-positive T-lymphocyte responses to various Mamu-A*01-bound peptides. The comparison of CD8-positive T-lymphocyte responses to 6 different epitopes in 10 Mamu-A*01-positive SIV-infected macaques during the first 12 weeks of infection shows that there is a strong CD8-positive T-lymphocyte response to Tat during the acute phase. The Mamu-A*01 Tat_{28–35} tetramer was initially constructed using an

SIV_{MAC251}-derived peptide (TTPESANL). This tetramer detected strong responses during the acute phase of SIV_{MAC239}-infected macaques, even though the corresponding SIV_{MAC239} sequence was STPESANL. Subsequent staining with the Tat_{28–35} STPESANL tetramer yielded identical results (data not shown).

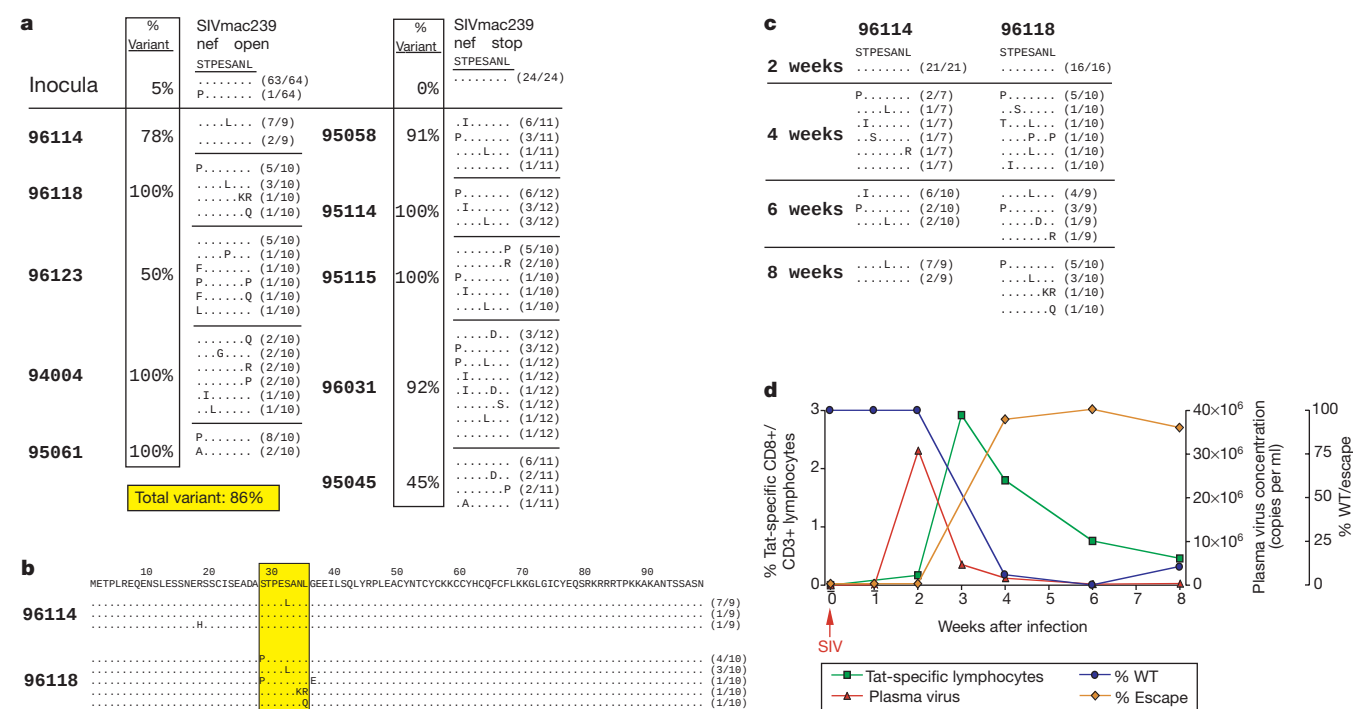


Figure 2 Variation is present in the SL8 epitope of all 10 Mamu-A*01-positive animals after acute SIV infection using analysis of clones isolated from plasma virus. **a**, Variation in the SL8 epitope in Mamu-A*01-positive animals infected with molecularly cloned SIV_{MAC239} Nef stop or SIV_{MAC239} Nef open. Limited variation is detectable in the inocula. The predicted amino-acid translation of a minimum of 9 clones isolated 6–8 weeks after infection is shown. The frequency of the epitope variant is shown to the right of the sequence. **b**, Little variation outside the SL8 epitope in the 5' exon of Tat in Mamu-A*01-positive animals. The entire 5' exon of Tat in two Mamu-A*01-positive animals is shown. Amino-acid substitutions accumulate primarily in the STPESANL epitope during the first

8 weeks of infection. **c**, Viral populations in the Mamu-A*01-restricted SL8 epitope evolve rapidly during the acute phase. Predicted amino-acid sequences of virus derived 2, 4, 6 and 8 weeks after infection in animals 96114 and 96118. By 4 weeks after infection, sequence variation was detectable within the epitope. **d**, Temporal relationships between viral load, tetramer levels and the percentage of wild-type and escaped virus. Values were averaged for animals 96114 and 96118, illustrating that loss of wild-type virus is coincident with the peak of tetramer levels and with the decline in plasma virus concentrations.

substitutions at anchor residues to have the most profound effect on peptide binding. Similarly, analyses of CTL lines generated from the peripheral blood mononuclear cells (PBMCs) of several Mamu-A*01-positive animals stimulated with the SL8 index peptide poorly recognized the new variant epitopes (Fig. 3). Notably, variant peptides with the P5 leucine substitution were the least efficient at sensitizing targets for CTL lysis, suggesting that this P5 mutation was probably interfering with T-cell receptor (TCR) recognition. Therefore, it seems probable that the new variants either reduced the amount of Tat-derived peptide/MHC class I complexes on the cell surface or reduced the ability of these complexes to be recognized by the T-cell receptor¹⁶.

To test the hypothesis that viruses with amino-acid replacements within the SL8 epitope are favoured by natural selection, we compared the number of synonymous nucleotide substitutions per synonymous site (d_s) and the number of non-synonymous nucleotide substitutions per non-synonymous site (d_N) in the

epitope and in the remainder of the sequence. In the SL8 epitope region of the virus from Mamu-A*01-positive animals, mean d_N was significantly higher than mean d_s both for comparisons between samples and the inoculum ($d_N = 5.7 \pm 0.4$; $d_s = 0.4 \pm 0.4$, $P < 0.001$) and in comparisons within samples ($d_N = 7.3 \pm 1.0$; $d_s = 0.7 \pm 0.7$, $P < 0.001$; see Supplementary Information, Table 1). Mean d_N values in the SL8 epitope from Mamu-A*01-positive animals were almost 60-fold greater than the corresponding values for Mamu-A*01-negative animals ($d_N = 5.7$ in Mamu-A*01-positive animals; $d_N = 0.1$ in Mamu-A*01-negative animals). As a pattern of $d_N > d_s$ is not expected under neutral evolution^{4,17,18}, this result strongly implies that amino-acid replacements in the SL8 epitope are favoured by positive darwinian selection.

Notably, the 5' exon of four of the eight Mamu-A*01-negative animals showed patterns of variation suggestive of escape from other Tat-specific cellular immune responses (see Supplementary Information, Figs 1 and 3). We therefore explored the possibility

Table 1 Peptide binding of mutant Tat epitopes to the rhesus MHC class I molecule Mamu-A*01

	IC ₅₀ *	Relative binding†		% Binding reduction‡
		(mean)	(s.d.)	
Tat _{28–35} epitope (index)	STPESANL	43	1	0
Variant	P.....	100	0.43	0.025
Variant	I.....	4,611	0.009	0.001
Variant	L.....	281	0.155	0.03
Variant	Q.....	13,175	0.003	0.0004
Variant	P.....	4,573	0.009	0.004
Variant	R.....	15,091	0.003	0.0012

* ¹²⁵I-labelled peptide: ATPYDINQML.

† Relative peptide binding is expressed relative to the binding affinity of the Tat_{28–35} STPESANL to the rhesus macaque MHC class I molecule Mamu-A*01.

‡ The percentage binding reduction is calculated as (1.0 – relative binding average) × 100.

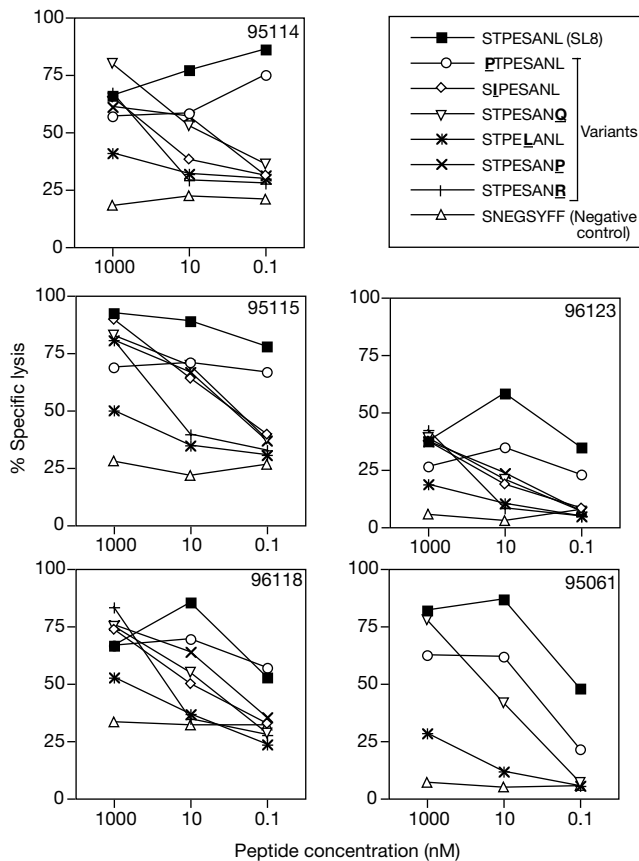


Figure 3 CTL analyses of CD8-positive T-lymphocytes stimulated with the SL8 peptide. Cell lines were stimulated with the index peptide (SL8) and autologous B-LCL. After 2 weeks in culture, these T-cell lines were used in CTL analyses at an effector:target (E:T) ratio of 25:1 with the index peptide and the variants. Three different dilutions of peptides were tested.

that animals with little evidence for selection in Tat should have higher plasma virus concentrations than animals with evidence for increased d_N in Tat. As the 18 animals in our cohort were originally part of a vaccine study¹⁹, we excluded the 8 Mamu-A*01-positive animals that had been vaccinated from this analysis. The two naive Mamu-A*01-positive and four of the naive Mamu-A*01-negative animals exhibited evidence of increased d_N peaks within the 5' exon of Tat, whereas four Mamu-A*01-negative naive animals revealed little evidence of increased d_N in Tat (see Supplementary Information, Fig. 4). Averaging the plasma virus concentrations of these two groups of animals showed a significant difference of at least one log ($P = 0.008$) between the plasma virus concentrations of animals with high and low d_N in Tat at 2 and 4 weeks after peak viraemia (see Supplementary Information, Fig. 5). Similarly, a significant inverse correlation was observed between peak d_N and viral load 2 weeks ($P = 0.007$), 4 weeks ($P = 0.008$) and 8 weeks ($P = 0.048$) after peak viraemia (Fig. 4; and data not shown). Of the four animals with low d_N in Tat, two rapidly progressed to simian AIDS and had SIV plasma virus concentrations in excess of 100×10^6 copies per ml within 6 months of infection. Therefore, animals with evidence of increased d_N in Tat may have controlled wild-type virus better than those with less selective pressure on Tat.

Vaccine-induced cellular immune responses against proteins expressed early in the viral life cycle may be better able to control HIV and SIV replication than responses directed against proteins that are expressed later in the viral life cycle. Viral escape from Tat-specific CD8-positive T lymphocytes occurred with kinetics similar to those seen during the emergence of drug-resistant mutants²⁰. In

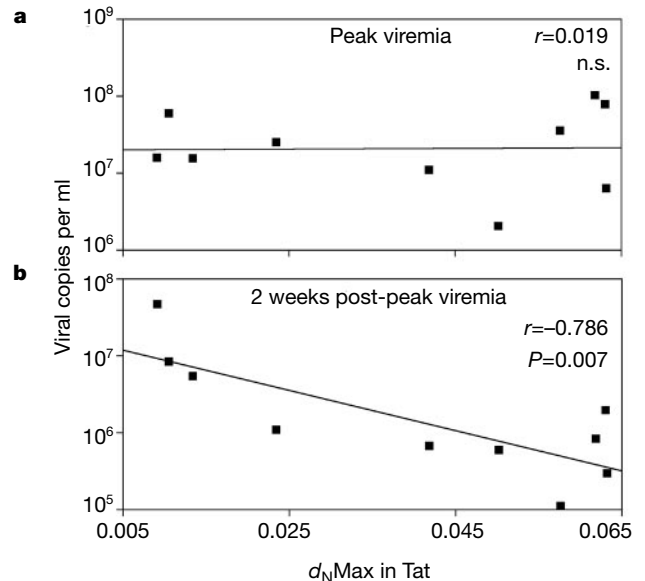


Figure 4 Inverse correlation between plasma viral concentrations and peak d_N in Tat. Peak d_N in Tat was plotted against plasma viral concentrations at peak viraemia and at 2 weeks after peak viraemia, revealing a significant inverse correlation between peak d_N and viral concentration at 2 weeks after peak viraemia.

five out of ten Mamu-A*01-positive animals, all clones isolated from plasma at 6–8 weeks after infection contained mutations in the Mamu-A*01-restricted SL8 epitope. This implies that Tat-specific CD8-positive T lymphocytes efficiently controlled replication of the original wild-type inoculum virus in these five animals. Responses directed against early proteins such as Tat may be particularly effective at controlling initial virus replication, as Tat and Rev are the only two viral proteins produced before Nef downregulates MHC class I molecules²¹. Tat-specific CTLs may, therefore, be potent inhibitors of early viral replication, whereas CTLs directed against peptides derived from other viral proteins may find few MHC class I/peptide complexes on the cell surface later in the course of the viral life cycle. The differences between the Gag and Tat-specific CTLs in their ability to exert selective pressure favouring viral escape are intriguing. Understanding the qualitative differences between these CTLs that account for these characteristics will be an important issue in the design of an effective HIV vaccine. Interestingly, vaccination of non-human primates with either Tat protein^{22,23} or recombinant viruses expressing Tat and Rev²⁴ have reduced virus replication. In these studies it is possible that Tat-specific CD8-positive T-lymphocyte responses were involved in control of viral replication. □

Methods

Tetramer analysis

Soluble tetrameric Mamu-A*01 MHC class I/SIV peptide complexes were constructed as described^{14,25}. Background tetramer staining of fresh, unstimulated PBMCs from naive Mamu-A*01-positive animals was routinely less than 0.08%.

Amplification of viral RNA from plasma and sequence detection

We obtained SIV plasma virus sequence as described⁷. The primers used to amplify complementary DNA encoding the Mamu-A*01 Tat epitope included SIV 6511-F (5'-TGATCCTCGTGTGCTAACTG-3') and 6900-R (5'-AGCAAGATGGCGATAAGCAG-3'). These primers were then used to isolate and sequence the cloned inserts. Seven overlapping PCR primer pairs (see Supplementary Information, Fig. 2b) were used to amplify cDNA spanning the entire SIV genome. The PCR products were directly sequenced from both cDNA strands. Overlapping sequence between the primers linked together sequences from the individual RT-PCR reactions. Sequence editing and finishing was performed using Auto Assembler v2.1 on a Macintosh. Nucleotide and predicted amino-acid sequences were aligned using MacVector 4.1 (Oxford Molecular).

Mamu-A*01 binding assay

We carried out quantitative assays for the binding of peptides to soluble Mamu-A*01 molecules on the basis of the inhibition of binding of a radiolabelled standard probe peptide to detergent-solubilized MHC molecules¹⁵. We used a position-1 C→A mutant of the SIV Gag 181–190 peptide (ATPYDINQML) as the radiolabelled probe. In the case of competitive assays, the concentration of peptide yielding 50% inhibition of the binding of the radiolabelled probe peptide was calculated. We initially tested peptides at one or two high doses. The half-maximal inhibitory concentration (IC₅₀) of peptides yielding positive inhibition were then determined in subsequent experiments, in which 2–6 further dilutions were tested, as necessary. Because, under the conditions we used, where [label] < [MHC] and IC₅₀ ≈ [MHC], the measured IC₅₀ values are reasonable approximations of the true K_d values. Each competitor peptide was tested in 2–4 completely independent experiments. As a positive control, in each experiment we tested the unlabelled version of the radiolabelled probe and measured its IC₅₀.

Generation of *in vitro* cultured CTL effector cells

We established CTL cultures from peripheral blood samples of SIV-infected rhesus macaques drawn in EDTA tubes, and cultured and assayed CTLs as described⁷.

Animals, viruses and infections

Rhesus macaques used in this study were identified as Mamu-A*01⁺ by PCR with sequence-specific primers (SSP) and direct sequencing as described¹⁶. All rhesus macaques used in this study were Mamu-A*01-positive, with the exception of animals 95003, 95112, 96020, 96081, 96093, 96072, 96104 and 96113. Rhesus macaques 95045, 96031, 95058, 96118, 96123, 95061, 96114 and 94004 were vaccinated with a DNA/Modified Vaccinia Ankara (MVA) regimen expressing the Gag_{181–189} peptide (CTPYDINQM)¹⁵. The Mamu-A*01-positive macaques 95114 and 95115 were not vaccinated before challenge. All rhesus macaques were infected intracutaneously with a molecularly cloned virus; SIV_{MAC239} (ref. 27) either Nef stop (95045, 96031, 95058, 95114, 95115, 95003 and 95112) or Nef open (99118, 96123, 95061, 96114, 94004, 96020, 96081, 96093, 96072, 96104 and 96113). Plasma viral concentrations were measured by branched DNA analysis (Chiron). The virus stock was amplified on rhesus PBMCs only. SIV-infected animals were cared for according to an experimental protocol approved by the University of Wisconsin Research Animal Resource Committee.

Statistical analysis of sequence and plasma virus concentration data

Numbers of synonymous nucleotide substitutions per synonymous site (d_s) and of non-synonymous nucleotide substitutions per non-synonymous site (d_N) were estimated as described²⁸. For the sample of viral sequences taken from a given animal, the means of d_s and d_N were computed for (1) all pairwise comparisons between each viral sequence and viral sequences sampled from the inoculum; and (2) all pairwise comparisons among viral sequences within the sample. These quantities were computed for the Tat_{28–35} epitope and the remainder of the 98-codon portion of Tat that was sequenced. To evaluate the statistical significance of the difference in peak and post-peak viraemia between macaques with high and low d_N in Tat, we compared the natural log (that is, ln) plasma virus concentrations among animals with high and low d_N . Taking ln greatly improved the fit of the data to the assumptions of the statistical models used (that is, normality, homoscedasticity in a multivariate test). We also used multiple analysis of variance (MANOVA) which is dependent on fewer assumptions than the repeated measures ANOVA.

Received 10 May; accepted 11 July 2000.

- Koup, R. A. *et al.* Temporal association of cellular immune responses with the initial control of viremia in primary human immunodeficiency virus type 1 syndrome. *J. Virol.* **68**, 4650–4655 (1994).
- Yasutomi, Y., Reimann, K., Lord, C., Miller, M. & Letvin, N. Simian immunodeficiency virus-specific CD8⁺ lymphocyte response in acutely infected rhesus monkeys. *J. Virol.* **67**, 1707–1711 (1993).
- Goulder, P. J. R. *et al.* Late escape from an immunodominant cytotoxic T-lymphocyte response associated with progression to AIDS. *Nature Med.* **3**, 212–217 (1997).
- Wolinsky, S. M. *et al.* Adaptive evolution of human immunodeficiency virus-type 1 during the natural course of infection. *Science* **272**, 537–542 (1996).
- Koenig, S. *et al.* Transfer of HIV-1-specific cytotoxic T lymphocytes to an AIDS patient leads to selection for mutant HIV variants and subsequent disease progression. *Nature Med.* **1**, 330–336 (1995).
- Mortara, L. *et al.* Selection of virus variants and emergence of virus escape mutants after immunization with an epitope vaccine. *J. Virol.* **72**, 1403–1410 (1998).
- Evans, D. T. *et al.* Virus-specific CTL responses select for amino acid variation in SIV Env and Nef. *Nature Med.* **5**, 1270–1276 (1999).
- Phillips, R. E. *et al.* Human immunodeficiency virus genetic variation that can escape cytotoxic T cell recognition. *Nature* **354**, 453–459 (1991).
- Coullin, I. *et al.* Impaired cytotoxic T lymphocyte recognition due to genetic variations in the main immunogenic region of the human immunodeficiency virus 1 NEF protein. *J. Exp. Med.* **180**, 1129–1134 (1994).
- Price, D. A. *et al.* Positive selection of HIV-1 cytotoxic T lymphocyte escape variants during primary infection. *Proc. Natl Acad. Sci. USA* **94**, 1890–1895 (1997).
- Borrow, P. *et al.* Antiviral pressure exerted by HIV-1-specific cytotoxic T lymphocytes (CTLs) during primary infection demonstrated by rapid selection of CTL escape virus. *Nature Med.* **3**, 205–211 (1997).
- Balter, M. Modest Briton stirs up storm with views on role of CTLs. *Science* **280**, 1860–1861 (1998).
- Meyerhans, A. *et al.* *In vivo* persistence of a HIV-1-encoded HLA-B27-restricted cytotoxic T lymphocyte epitope despite specific *in vitro* reactivity. *Eur. J. Immunol.* **21**, 2637–2640 (1991).
- Allen, T. M. *et al.* CD8⁺ lymphocytes from SIV-infected rhesus macaques recognize 27 different peptides bound by a single MHC class I molecule: Implications for vaccine design and testing. *J. Virol.* (in the press).

- Allen, T. M. *et al.* Characterization of the peptide-binding motif of a rhesus MHC class I molecule (Mamu-A*01) that binds an immunodominant CTL epitope from SIV. *J. Immunol.* **160**, 6062–6071 (1998).
- Price, G. E., Ou, R., Jiang, H., Huang, L. & Moskopidhis, D. Viral escape by selection of cytotoxic T cell-resistant variants in influenza A virus pneumonia. *J. Exp. Med.* **191**, 1853–1867 (2000).
- Kimura, M. Preponderance of synonymous changes as evidence for the neutral theory of molecular evolution. *Nature* **267**, 275–276 (1977).
- Hughes, A. L. & Nei, M. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**, 167–170 (1988).
- Allen, T. M. *et al.* Induction of AIDS virus-specific CTL activity in fresh, unstimulated PBL from rhesus macaques vaccinated with a DNA prime/MVA boost regimen. *J. Immunol.* **164**, 4968–4978 (2000).
- Coffin, J. M. HIV population dynamics *in vivo*: Implications for genetic variation, pathogenesis, and therapy. *Science* **267**, 483–489 (1995).
- Collins, K. L., Chen, B. K., Kalams, S. A., Walker, B. D. & Baltimore, D. HIV-1 Nef protein protects infected primary cells against killing by cytotoxic T lymphocytes. *Nature* **391**, 397–401 (1998).
- Cafaro, A. *et al.* Control of SHIV-89.6P-infection of cynomolgus monkeys by HIV-1 Tat protein vaccine. *Nature Med.* **5**, 643–650 (1999).
- Pauza, C. D. *et al.* Vaccination with Tat toxoid attenuates disease in simian/HIV-challenged macaques. *Proc. Natl Acad. Sci. USA* **97**, 3515–3519 (2000).
- Osterhaus, A. D. M. E. *et al.* Vaccination with Rev and Tat against AIDS. *Vaccine* **17**, 2713–2714 (1999).
- Altman, J. D. *et al.* Phenotypic analysis of antigen-specific T lymphocytes. *Science* **274**, 94–96 (1996).
- Knapp, L. A., Lehmann, E., Piekarczyk, M. S., Urvater, J. A. & Watkins, D. I. A high frequency of Mamu-A*01 in the rhesus macaque detected by polymerase chain reaction with sequence-specific primers and direct sequencing. *Tissue Antigens* **50**, 657–661 (1997).
- Regier, D. A. & Desrosiers, R. C. The complete nucleotide sequence of a pathogenic molecular clone of simian immunodeficiency virus. *AIDS Res. Hum. Retroviruses* **6**, 1221–1231 (1990).
- Nei, M. & Gojobori, T. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**, 418–426 (1986).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank L. Smith and B. Becker for preparation of this manuscript and C. D. Pauza and the Immunology and Virology Core Laboratory for infection with molecularly cloned SIV_{MAC239} nef stop and monitoring of macaques. This work was supported by the NIAID, NCRR and the The James B. Pendleton Charitable Trust. D.I.W. is an Elizabeth Glaser scientist.

Correspondence and requests for materials should be addressed to D.I.W. (e-mail: watkins@primate.wisc.edu).

Superoxide dismutase as a target for the selective killing of cancer cells

Peng Huang*, Li Feng*, Elizabeth A. Oldham*, Michael J. Keating† & William Plunkett*

*Department of Experimental Therapeutics, †Department of Leukaemia, The University of Texas M. D. Anderson Cancer Center, 1515 Holcombe Boulevard, Houston, Texas 77030, USA

Superoxide dismutases (SOD) are essential enzymes that eliminate superoxide radical (O₂^{•−}) and thus protect cells from damage induced by free radicals^{1–3}. The active O₂^{•−} production and low SOD activity in cancer cells^{3–7} may render the malignant cells highly dependent on SOD for survival and sensitive to inhibition of SOD. Here we report that certain oestrogen derivatives selectively kill human leukaemia cells but not normal lymphocytes. Using complementary DNA microarray and biochemical approaches, we identify SOD as a target of this drug action and show that chemical modifications at the 2-carbon (2-OH, 2-OCH₃) of the derivatives are essential for SOD inhibition and for apoptosis induction. Inhibition of SOD causes accumulation of cellular O₂^{•−} and leads to free-radical-mediated damage to mitochondrial membranes, the release of cytochrome *c* from mitochondria and apoptosis of the cancer cells. Our results indicate that targeting SOD may be a promising approach to the selective killing of cancer cells, and that mechanism-based combinations of

SOD inhibitors with free-radical-producing agents may have clinical applications.

2-Methoxyoestradiol (2-ME), an oestrogen derivative that cannot bind the oestrogen receptor⁸, induces accumulation of p53 protein⁹. We initially used 2-ME to increase p53 protein in an attempt to enhance cellular response to other anticancer agents. Unexpectedly, 2-ME itself showed a potent activity against human leukaemia cells with different p53 genotypes. Typical apoptotic morphology and nucleosomal DNA fragmentation were observed in ML-1 (wild-type p53) and HL-60 (p53⁻) cells treated with 1 μ M 2-ME (Fig. 1a).

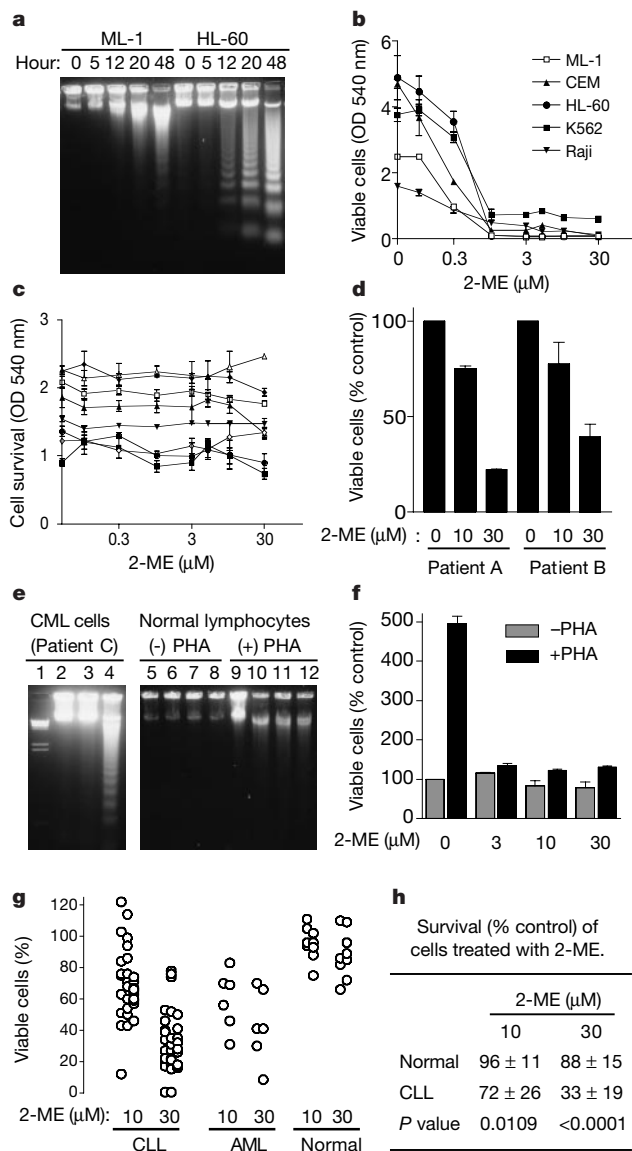


Figure 1 Selective cytotoxicity of 2-ME against human leukaemia cells. **a**, DNA fragmentation in ML-1 and HL-60 cells incubated with 1 μ M 2-ME. **b**, MTT assay of leukaemia cells incubated with 2-ME (72 h). **c**, Effect of 2-ME on normal lymphocytes from healthy donors (MTT assay, 72 h). **d**, Leukaemia cells from 2 CLL patients were incubated with 2-ME for 72 h followed by MTT assay. **e**, DNA fragmentation in CML cells and normal lymphocytes. Lane 1, markers; lanes 2–4, CML cells incubated with medium, 1 μ M ara-C and 10 μ M 2-ME (48 h); lanes 5–8, lymphocytes incubated with 0, 3, 10 and 30 μ M 2-ME for 48 h; lanes 9–12, PHA-stimulated (5 μ g ml⁻¹) lymphocytes incubated with 0, 3, 10 and 30 μ M 2-ME for 48 h. **f**, Effect of 2-ME on normal and PHA-stimulated lymphocytes (MTT assay, 72 h). **g**, Effect of 2-ME on primary leukaemia cells and normal lymphocytes (MTT assay). **h**, Comparison of 2-ME activity in CLL cells ($n = 31$) and normal lymphocytes ($n = 9$).

The MTT (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide) assay further showed that cells with wild-type p53 (ML-1), mutant p53 (CEM and Raji) and no p53 (K562 and HL-60) were similarly sensitive to 2-ME (Fig. 1b). The half-maximal inhibitory concentration (IC₅₀) was less than 1 μ M for all five cell lines.

When normal lymphocytes from eight healthy donors were incubated with 2-ME, there was no apparent reduction in cell survival (Fig. 1c), indicating that 2-ME may have a selective anti-leukaemia activity. Because normal lymphocytes were quiescent and might not be comparable to proliferating leukaemia cells, we used

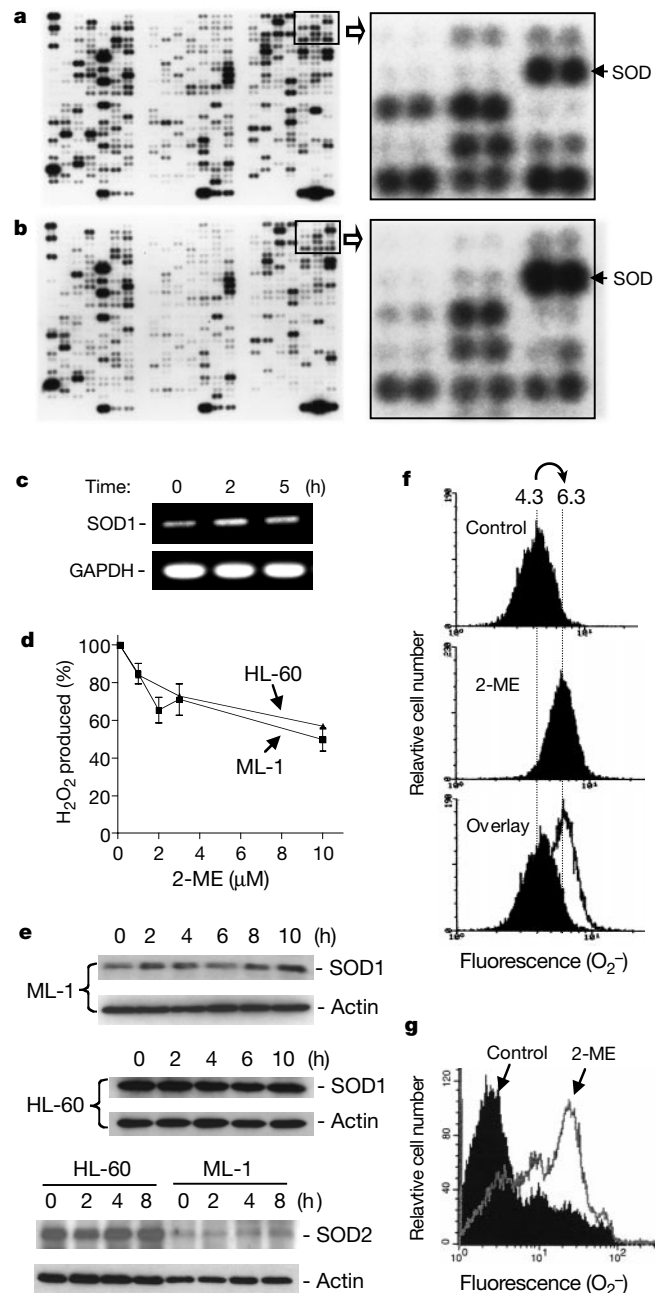


Figure 2 Effect of 2-ME on SOD expression and free-radical metabolism in leukaemia cells. **a**, **b**, Gene expression profiles in untreated and 2-ME-treated (1 μ M, 5 h) ML-1 cells, respectively. The boxed areas containing CuZnSOD are enlarged and shown on the right. **c**, RT-PCR analysis of SOD1 mRNA in 2-ME-treated cells. **d**, Effect of 2-ME (2 h) on H₂O₂ production in ML-1 and HL-60 cells. **e**, Effect of 2-ME (2 μ M) on SOD1 and SOD2 protein expression in ML-1 and HL-60 cells. **f**, O₂⁻ accumulation in ML-1 cells treated with 1 μ M 2-ME for 5 h. Dotted lines, mean fluorescence intensity. **g**, Accumulation of O₂⁻ in primary CLL cells treated with 10 μ M 2-ME for 24 h.

quiescent leukaemia cells isolated from patients with chronic lymphocytic leukaemia (CLL) for comparison with normal lymphocytes. As shown in Fig. 1d, 2-ME substantially reduced the survival of CLL cells. This compound also induced apoptosis in leukaemia cells from a CML (chronic myelogenous leukaemia) patient in relapsed blast crisis (Fig. 1e, lane 4) that were resistant to ara-C *in vitro* (lane 3). Evaluation of the activity of 2-ME in primary leukaemia cells from thirty-one CLL patients, three CML patients, six acute myelogenous leukaemia (AML) patients and two acute lymphocytic leukaemia (ALL) patients revealed substantial activity in most cases, although the degree of cytotoxicity varied between samples (Fig. 1g). Overall, 30 μ M 2-ME reduced cell survival to 32.7% in CLL cells, 54.7% in CML cells, 49.6% in AML cells and 56.5% in ALL cells. 2-ME was significantly more toxic to CLL cells than to normal lymphocytes at both 10 and 30 μ M ($P < 0.0001$, Fig. 1h).

To test the effect of 2-ME on normal lymphocytes induced to re-enter the proliferating cell cycle, we stimulated lymphocytes with phytohaemagglutinin (PHA) and then exposed them to 2-ME. PHA (5 μ g ml⁻¹, 48 h) induced significant proliferation, as measured by [³H]thymidine incorporation (6,903 d.p.m. $\rightarrow$ 99,875 d.p.m.) and by flow cytometry analysis (S phase: 0% $\rightarrow$ 17.1%; G2/M phases: 0% $\rightarrow$ 9.5%; $n = 3$). When the lymphocytes were incubated with 2-ME (0–30 μ M) for three more days, no nucleosomal DNA fragmentation was detected (Fig. 1e, lanes 5–12). A small

amount of DNA smearing was seen in all PHA-stimulated cells (lanes 9–12), which may indicate some spontaneous turnover of the cells during the five-day incubation. MTT assay revealed that 2-ME inhibited the proliferation of the PHA-treated lymphocytes, which nevertheless remained viable during the three-day drug incubation (Fig. 1f).

We used a cDNA microarray assay (Atlas Human cDNA Expression Array, ClonTech) to search for candidate molecules involved in the cellular responses to 2-ME. The gene in the second row of the last column in Fig. 2b (in doublets) showed a significant increase in messenger RNA expression 5 h after incubation with 2-ME (compare with Fig. 2a). This molecule was identified as CuZnSOD (SOD1; GenBank accession no. k00065). Quantification by phosphorimager revealed that expression of CuZnSOD in 2-ME-treated cells was 236% of control, after the signal was normalized by three internal standards (see Methods). This increase was confirmed by a reverse transcriptase polymerase chain reaction (RT-PCR) assay (Fig. 2c). A similar increase in CuZnSOD mRNA expression was also observed in HL-60 cells (not shown).

To investigate the mechanism responsible for the increase in SOD mRNA expression, we tested two possibilities. (1) 2-ME might cause an increase in cellular O₂⁻ production and thus induce SOD expression in response to the free-radical stress. This would lead to an increase in H₂O₂ production. (2) 2-ME might inhibit SOD enzyme activity and cause a feedback upregulation of SOD expression. This would lead to a decrease in H₂O₂ generation. In fact, 2-ME caused a concentration-dependent decrease in H₂O₂ in ML-1 and HL-60 cells (Fig. 2d). Immunoblot analysis showed that the reduced H₂O₂ production was not due to a loss of CuZnSOD or MnSOD protein (Fig. 2e). Rather, a moderate increase (74%) in SOD1 protein was seen in ML-1 cells at 10 h. Thus, the increase in SOD mRNA expression probably reflects a cellular response to O₂⁻ accumulation due to SOD inhibition by 2-ME. To confirm the accumulation of O₂⁻ in 2-ME-treated cells, we used hydroethidine, a compound specifically converted by O₂⁻ to highly fluorescent ethidium^{10,11}, to measure the cellular O₂⁻ content by flow cytometry analysis. Figure 2f shows that treatment of ML-1 cells with 1 μ M 2-ME for 5 h caused an increase in O₂⁻ from 4.3 to 6.3 arbitrary units. Incubation of CLL cells with 30 μ M 2-ME also led to a substantial O₂⁻ increase (226 $\pm$ 37 %, Fig. 2g).

We used an *in vitro* assay for SOD activity¹² to test the direct effect of 2-ME on purified CuZnSOD and MnSOD. The xanthine/xanthine oxidase system was used to generate O₂⁻, which reacted with XTT to produce a reddish product (470 nm). Addition of CuZnSOD to the reaction caused dismutation of O₂⁻ and reduced the absorbance at 470 nm in a concentration-dependent manner (Fig. 3a). 2-ME prevented the reduction of absorbance at 470 nm in the presence of SOD (Fig. 3b), indicating that 2-ME inhibited SOD activity. The inhibition of human CuZnSOD, bovine CuZnSOD and the *Escherichia coli* MnSOD by 2-ME was concentration dependent (IC₅₀ $\approx$ 20 μ M, Fig. 3c). Furthermore, 2-ME (1–100 μ M) had little effect on xanthine oxidase (Fig. 3d), DNA polymerase- α or alkaline phosphatase (Fig. 3e), indicating that inhibition of SOD by 2-ME is probably specific. To test whether there was a direct interaction between 2-ME and SOD protein, extracts of HL-60 cells incubated with [³H]2-ME (1 μ M, 0.3 μ Ci ml⁻¹, 5 h) were immunoprecipitated with antibodies against human CuZnSOD, MnSOD, β -actin or histone (common epitope). The radioactivity associated with SOD1 or SOD2 was significantly higher than that associated with β -actin or histone (Fig. 3f), indicating that 2-ME may bind preferentially to SODs in intact cells.

Four structurally related oestrogen derivatives were compared with 2-ME to investigate the chemical basis for inhibition of SOD. As shown in Fig. 4a, 2-ME (no. 1), 2-hydroxyoestradiol (no. 4) and 2-methoxyoestrone (no. 5), which have an -OH or -OCH₃ group at the 2-carbon, substantially inhibited SOD activity and induced DNA fragmentation. In contrast, 17 β -oestradiol (no. 2) and oes-

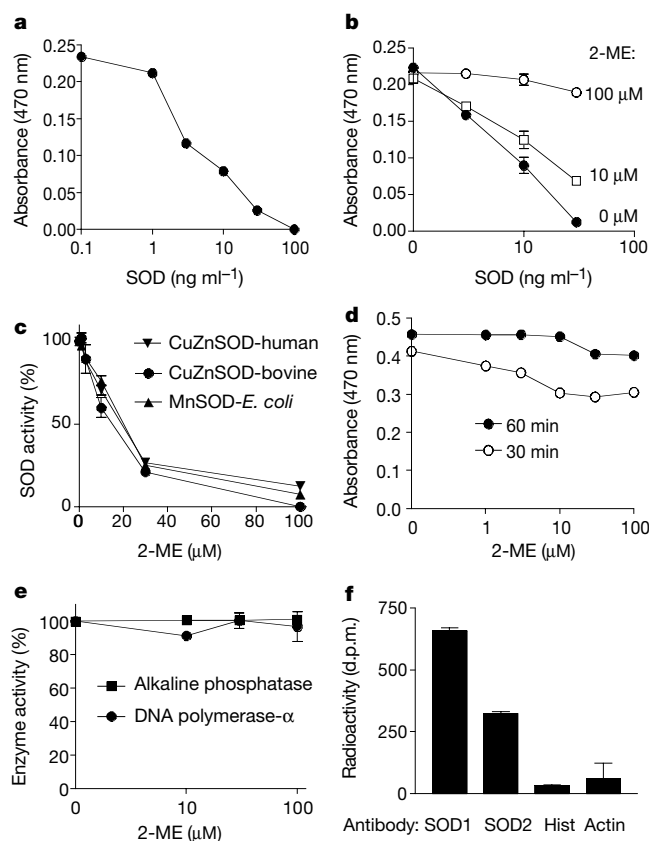


Figure 3 Inhibition of SOD by 2-ME. **a**, *In vitro* assay of SOD1 activity (see Methods). **b**, Inhibition of SOD1 by 2-ME *in vitro*. **c**, Effect of 2-ME on human and bovine CuZnSOD and *E. coli* MnSOD. **d**, Effect of 2-ME on xanthine oxidase. **e**, Effect of 2-ME on human DNA polymerase- α ²⁷ and bovine alkaline phosphatase (measured by removal of 5'-phosphate from [¹⁴C]GMP). **f**, Binding of 2-ME to SOD. HL-60 cells were incubated with [³H]2-ME (1 μ M, 0.3 μ Ci ml⁻¹, 5 h), washed and sonicated in PBS containing protease inhibitors. The protein extracts were precipitated with antibodies against CuZnSOD, MnSOD, β -actin or histone and protein-G/sepharose conjugate. Radioactivity associated with the washed pellets was determined by liquid scintillation counting.

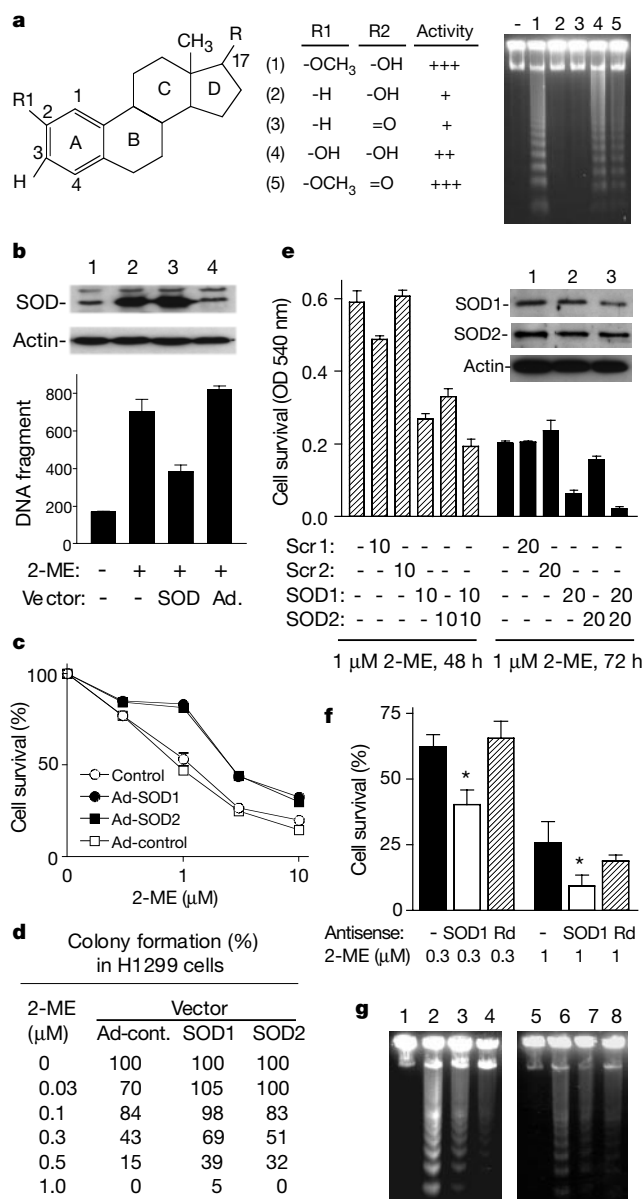


Figure 4 Relationship between SOD activity and apoptosis in 2-ME-treated cells. **a**, Structures of oestrogen derivatives, inhibition of CuZnSOD and induction of apoptosis in HL-60 cells. Degree of SOD inhibition: -, <25%; +, 25–50%; ++, 50–75%; +++, 75–100%. Lane -, control; lanes 1–5, cells treated with the respective numbered compounds. **b**, Effect of SOD1 overexpression on 2-ME-induced apoptosis. Lane 1, control A2008 cells; lanes 2, 3, transduction with Ad.CuZnSOD for 24 and 48 h, respectively; lane 4, control vector (48 h). 2-ME (10 μM) was added 24 h after transduction and incubated for another 48 h before DNA fragmentation assay. **c**, **d**, Effect of ectopic expression of SOD1 or SOD2 on the survival of A2008 cells (MTT assay) and H1299 cells (colony formation). **e**, SOD antisense S-oligonucleotides enhanced 2-ME activity. A2008 cells were incubated with S-oligonucleotides against SOD1 (5'-ACGCACACGGCCTTCGTGCCATAACT) and SOD2 (5'-GCACACTGCCCGGCTCAA-CATGCTG) or the respective scrambled S-oligonucleotides (Scr). SOD protein levels were assayed at 24 h. Lane 1, control; lane 2, 10 μM each of the scrambled S-oligonucleotides; lane 3, 10 μM each of anti-SOD1 and anti-SOD2 S-oligonucleotides. 2-ME was added at 24 h and incubated for another 48–72 h followed by MTT assay. **f**, Effect of anti-SOD1 or random (Rd) S-oligonucleotide (10 μM) on the survival of 2-ME-treated H1299 cells (MTT, 72 h). **g**, Effect of antioxidants on 2-ME-induced apoptosis. Lanes 1 and 5, control HL-60 cells; lanes 2–4, cells treated with 1 μM 2-ME plus 0, 0.03 and 0.1 mM ambroxol; lanes 6–8, cells treated with 1 μM 2-ME plus 0, 3 and 5 mM N-acetylcysteine.

trone (no. 3), which lack the 2-carbon modification, showed minimal activity against SOD and caused little DNA fragmentation. Thus, a 2-OH or 2-OCH₃ modification is important for inhibiting SOD and inducing apoptosis.

If inhibition of SOD were critical for the cytotoxic action of 2-ME, one would expect that a change in SOD expression would alter cellular sensitivity to 2-ME. When human ovarian cancer cells (A2008) were infected with an adenoviral vector containing full-length CuZnSOD sequence (Ad.CuZnSOD)¹³, substantial increases in SOD1 protein were detected at 24 and 48 h (Fig. 4b, lanes 2 and 3). This overexpression led to a decrease in 2-ME-induced apoptosis, as shown by a reduction in nucleosomal DNA fragmentation (Fig. 4b). In contrast, infection with the control viral vector did not affect SOD expression (lane 4) or drug sensitivity. An MTT assay also revealed partial protection of the 2-ME-treated cells by transduction with Ad.CuZnSOD or Ad.MnSOD (Fig. 4c). Colony formation assay in another cell line (H1299) confirmed this protective effect (Fig. 4d).

We used synthetic antisense S-oligonucleotides against SOD1 and

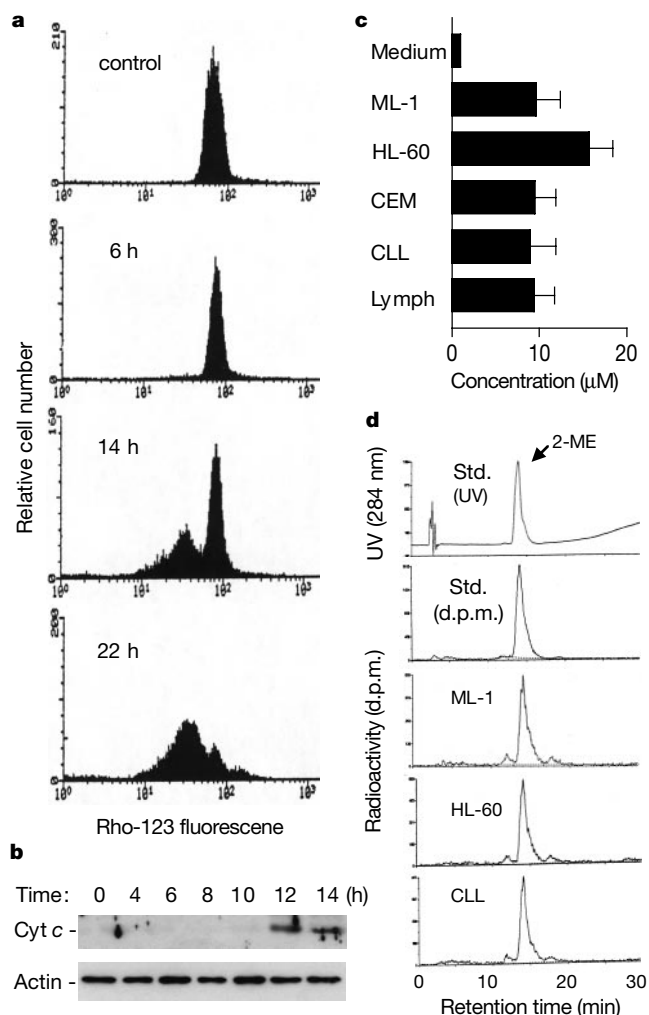


Figure 5 Cellular uptake and metabolism of 2-ME and its effect on mitochondrial membrane integrity. **a**, 2-ME-induced loss of mitochondrial membrane integrity. HL-60 cells were incubated with 2 μM 2-ME for up to 22 h. The integrity of mitochondrial membranes was measured using rhodamine-123 as described^{14,15}. **b**, Release of cytochrome c to cytosol in cells treated with 2 μM 2-ME. **c**, Accumulation of 2-ME in leukaemia cells and normal lymphocytes (see Methods). **d**, HPLC analysis of extracts from cells incubated with [3H]2-ME (0.5 μCi ml⁻¹, 5 h). The top two panels are chromatograms of [3H]2-ME standard monitored by ultraviolet (UV) and liquid scintillation counting, respectively. The lower three panels show the profiles of radioactivity in extracts of ML-1, HL-60 and CLL cells.

SOD2 to suppress SOD expression, and their effect on cellular sensitivity to 2-ME was then evaluated. As shown in Fig. 4e, incubation of A2008 cells with 10 μM each of the anti-SOD S-oligonucleotides for 48 h led to a moderate decrease in SOD1 and SOD2 protein. This was associated with an increase in 2-ME-induced cytotoxicity. A longer incubation (72 h) with 20 μM anti-SOD S-oligonucleotides resulted in greater enhancement of 2-ME activity. Treatment with scrambled S-oligonucleotides did not alter the SOD protein levels or the cellular sensitivity to 2-ME (Fig. 4e). Significant enhancement of the activity of 2-ME (0.3 μM , $P = 0.0059$; 1 μM , $P = 0.0452$) by anti-SOD1 S-oligonucleotides was also observed in H1299 cells (Fig. 4f). The role of SOD inhibition by 2-ME in causing apoptosis was supported by experiments with antioxidants. When HL-60 cells were incubated with 1 μM 2-ME in the presence of the O_2^- scavenger ambroxol or antioxidant *N*-acetylcysteine, the 2-ME-induced apoptosis was significantly reduced (Fig. 4g). The protective effect of *N*-acetylcysteine was also observed in CLL cells using the MTT assay (not shown).

Because mitochondria are a major source of superoxide (O_2^-) production, we reasoned that inhibition of SOD by 2-ME might cause mitochondrial damage owing to free-radical attack on the membrane phospholipids. To test this possibility, the integrity of mitochondrial membranes of 2-ME-treated cells was examined by measuring their ability to retain rhodamine-123, a fluorescent dye used to indicate the loss of mitochondrial transmembrane potential^{14,15}. Under the experimental conditions (rhodamine-123, 5 $\mu\text{g ml}^{-1}$, 30 min), the control HL-60 cells retained 70–100 arbitrary units of fluorescence (Fig. 5a). During the first 6 h of 2-ME incubation, no obvious loss of dye retention was observed. However, significant loss of membrane integrity was seen after prolonged exposure to 2-ME (14 h), as shown by a shift of the fluorescence intensity to much lower levels (20–40 units). By 22 h, almost all the cells had lost their ability to retain rhodamine-123 in the mitochondria.

Cytochrome *c* is normally located in mitochondria and its release to the cytosol triggers apoptosis¹⁶. Thus, we tested the possibility that mitochondrial membrane damage by 2-ME might result in the release of cytochrome *c*. As shown in Fig. 5b, no cytochrome *c* was present in cytosol during the first 10 h. However, significant amounts of cytosolic cytochrome *c* were detected at 12 and 14 h. This was in agreement with the time course of mitochondrial membrane damage. In separate experiments, 2-ME-induced apoptosis occurred in the presence of cycloheximide (5–50 $\mu\text{g ml}^{-1}$) or actinomycin D (1 $\mu\text{g ml}^{-1}$), indicating that induction of apoptosis by 2-ME may not require the synthesis of new RNA or protein. This is consistent with the mechanism of action of 2-ME in causing free-radical-mediated damage to mitochondrial membrane and release of cytochrome *c*, which then activated the apoptotic cascade.

In the leukaemia cell lines tested, 2-ME induced apoptosis at concentrations of 1–10 μM , whereas the IC_{50} for SOD inhibition *in vitro* was about 20 μM . This apparent discrepancy was probably due to a concentrating uptake of 2-ME by the cells. To confirm this, we incubated leukaemia cell lines, primary CLL cells and normal lymphocytes with 1 μM [^3H]2-ME, and determined the intracellular 2-ME concentrations. As shown in Fig. 5c, the drug was concentrated about tenfold in most cases during the 5-h incubation. High performance liquid chromatography (HPLC) analysis of extracts from ML-1, HL-60 and CLL cells incubated with [^3H]2-ME revealed no significant drug metabolites (Fig. 5d). 2-ME appeared as the only major peak (14.1 min) in all samples; there were two barely detectable peaks at 11.8 and 17.6 min.

SODs catalyse the dismutation of O_2^- , an essential step in eliminating the toxic free radical. This critical function makes SOD an attractive target for pharmacological intervention. Although several small molecules, including cyanide ion (CN^-), hydroxyl ion (OH^-) and azide ion (N_3^-), were found to inhibit SOD by competing with O_2^- at the catalytic site¹⁷, these chemicals are

highly toxic and their potential for cancer therapy is limited. Here we have shown that certain oestrogen derivatives inhibit SODs and selectively induce apoptosis in human leukaemia cells. Inhibition of SOD leads to free-radical-mediated damage to mitochondrial membranes and the release of cytochrome *c*. This is probably a major pathway of 2-ME-induced apoptosis, although other mechanisms such as inhibition of tubulin polymerization and angiogenesis may also contribute to its activity^{8,18,19}. Because O_2^- may also give rise to hydroxyl radicals and cause DNA damage^{20,21}, this may explain 2-ME-induced accumulation of p53. The correlation between the ability of oestrogen derivatives to inhibit SOD and their ability to cause apoptosis indicates that inhibition of SOD may be an important mechanism by which 2-ME kills cancer cells. Alterations in the sensitivity of cells to 2-ME by modulating SOD expression with viral vectors or antisense oligonucleotides further support this conclusion. However, the role of O_2^- in causing apoptosis may vary among different cell types²², and thus inhibition of SOD should not be considered as a general approach to cancer therapy. Because inhibition of SOD impairs the ability of cancer cells to cope with O_2^- radical stress, it may also provide a mechanism for the design of new strategies for cancer treatment. For example, combinations of the SOD inhibitor (2-ME) with modalities that produce free radicals may result in synergistic anticancer activity. Ionizing radiation and certain drugs such as anthracyclines and bleomycin produce free radicals^{23–25} and thus are the logical choices for combination with 2-ME. Such mechanism-based strategies may have potential clinical applications and warrant further investigation. □

Methods

Assay of SOD activity

We used a spectrophotometric assay¹² to determine the effect of 2-ME on SOD activity *in vitro* with purified enzymes (bovine CuZnSOD from Boehringer Mannheim; human CuZnSOD and *E. coli* MnSOD from Sigma). The reactions contained 2.5 ml of 50 mM Na_2CO_3 , 0.1 ml of 3 mM xanthine, 0.1 ml of 3 mM EDTA and 0.1 ml of 0.8 mM XTT (3'-(1-[phenylamino-carbonyl]-3,4-tetrazolium)-bis(4-methoxy-6-nitro)benzenesulphonic acid hydrate), and the indicated concentrations of SOD and 2-ME. Xanthine oxidase (0.1 ml, 64 mU ml^{-1}) was added to start the reaction. After incubation at 24 °C for 30 min, the absorbance at 470 nm was measured. The relative activity (RA) of SOD was calculated as:

$$\text{RA} = [1 - (c - b)/(a - b)] \times 100\%$$

where *a* is the absorbance of reaction without SOD, *b* is the absorbance of reaction with SOD but without 2-ME and *c* is the absorbance of reaction with SOD in the presence of 2-ME. All values were subtracted by the background reading of the blank. In separate experiments using nitro blue tetrazolium or cytochrome *c* as substrates for SOD assay, 2-ME and DMSO (as a solvent for 2-ME) interfered with the colorimetric reaction. These artefacts preclude the use of NBT or cytochrome *c* as substrates for measuring the effect of 2-ME on SOD activity. Neither 2-ME nor DMSO affects the colorimetric reaction of XTT.

Quantification of cellular O_2^- and H_2O_2 production

We adapted the use of hydroethidine to detect O_2^- in tissue sections¹⁰ for quantification of intracellular O_2^- by flow cytometry analysis. This assay is based on the chemical properties of hydroethidine, a weak blue fluorescent dye which is selectively converted by O_2^- to ethidium with a bright red fluorescence^{10,11}. The control and drug-treated cells (1.5×10^6 per sample) were incubated with hydroethidine (20 ng ml^{-1}) for 60 min, washed twice with 3 ml PBS, resuspended in 2 ml PBS and then analysed within 1 h by flow cytometry, using the red laser channel. A microassay was used to measure the effect of 2-ME on H_2O_2 generation in cell culture as described²⁶.

Assay of gene expression by cDNA microarray analysis

We used the Atlas Human cDNA Expression Array I and the Atlas Human Cancer cDNA Expression Array (Clontech). We converted 1 μg of mRNA isolated from the control or drug-treated cells to radioactive cDNA by reverse transcription in the presence of [$\alpha\text{-}^{32}\text{P}$]-dATP. The ^{32}P -labelled cDNA was then denatured and hybridized to the cDNA expression arrays as recommended by the manufacturer. The radioactivity on the membranes was quantified by a phosphorimager. We calculated the change in gene expression after 2-ME treatment as the percentage of the untreated cells, using three of the internal controls (ubiquitin, GAPDH and the *M*, 23,000 highly basic protein) recommended by the manufacturer for normalization to ensure the comparability of the control and drug-treated samples.

RT-PCR analysis

The expression of SOD RNA was also measured by RT-PCR. RNA (1 μg) isolated from the control or 2-ME-treated cells was first converted to cDNA by reverse transcription using

SuperScript II reverse transcriptase (GibcoBRL) and the anchored oligo-dT primer set (Genosys). The cDNA was then amplified by PCR. The SOD1 primers were from Genosys (forwards, 5'-ACGAAGGCCGTGTGCGTCTGAA; backwards, 5'-ACCACAAGC-CAAACGACTTCCAGC). The reaction was run at 94 °C (1 min), 60 °C (1 min), 72 °C (1 min) for 20 cycles, which was within the linear reaction window. GAPDH was also measured by RT-PCR from the same RNA samples and used as an internal control (GAPDH primers: forwards, 5'-CCATCAATGACCCCTTCAATTGACC; backwards, 5'-GAAGGCCATGCCAGTGAGCTTCC).

Assays of cellular accumulation and metabolism of 2-ME

To determine cellular uptake of 2-ME, cells were incubated with 1 μM [^3H]2-ME (0.5 $\mu\text{Ci ml}^{-1}$) for 5 h and washed twice with cold PBS. Radioactivity associated with the cell pellets or the culture medium was quantified by liquid scintillation counting. 2-ME concentrations were calculated based on the specific radioactivity of [^3H]2-ME, cell number and the mean cell volume as measured by a Coulter Counter equipped with a 256-Channelizer. We used HPLC to analyse potential cellular metabolites of 2-ME. Cells were incubated with [^3H]2-ME (0.5 $\mu\text{Ci ml}^{-1}$) for 5 h, washed twice with PBS and then extracted with 50% methanol. The extracts were analysed by HPLC with an ultraviolet detector (284 nm) and an online liquid scintillation counter, using a $\mu\text{Bondapak C}_{18}$ reversed-phase column and the following running conditions: 1 ml min $^{-1}$; 0–5 min, 100% buffer A (water:acetonitrile:acetic acid, 59:40:1); 5–26 min, a linear gradient of 0–100% buffer B (water:acetonitrile:acetic acid, 39:60:1); 26–30 min, 100% buffer B.

Received 25 April; accepted 20 June 2000.

- Halliwell, B. & Gutteridge, J. M. C. *Free Radicals in Biology and Medicine* (Oxford Univ. Press, New York, 1985).
- Fridovich, I. Superoxide radical and superoxide dismutase. *Annu. Rev. Biochem.* **64**, 97–112 (1995).
- Oberley, L. W. & Buettner, G. R. Role of superoxide dismutase in cancer: a review. *Cancer Res* **39**, 1141–1149 (1979).
- Dionisi, O., Galeotti, T., Terranova, T. & Azzi, A. Superoxide radicals and hydrogen peroxide formation in mitochondria from normal and neoplastic tissues. *Biochim. Biophys. Acta* **403**, 292–300 (1975).
- Marklund, S. L., Westman, N. G., Lundgren, E. & Roos, G. Copper- and zinc-containing superoxide dismutase, manganese-containing superoxide dismutase, catalase, and glutathione peroxidase in normal and neoplastic human cell lines and normal human tissues. *Cancer Res.* **42**, 1955–1961 (1982).
- Van Driel, B. E. & Van Noorden, C. J. Oxygen insensitivity of the histochemical assay of glucose-6-phosphate dehydrogenase activity for the discrimination between nonmalignant and malignant cells. *J. Histochem. Cytochem.* **47**, 575–582 (1999).
- Shulyakovskaya, T., Sumegi, L. & Gal, D. In vivo experimental studies on the role of free radicals in photodynamic therapy. I. Measurement of the steady state concentration of free radicals in tumor tissues of mice. *Biochem. Biophys. Res. Commun.* **95**, 581–587 (1993).
- Cushman, M., He, H.-M., Katzenellenbogen, J. A., Lin, C. M. & Hamel, E. Synthesis, antitubulin and antimetabolic activity, and cytotoxicity of analogs of 2-methoxyestradiol, an endogenous mammalian metabolite of estradiol that inhibits tubulin polymerization by binding to the colchicine binding site. *J. Med. Chem.* **38**, 2041–2049 (1995).
- Mukhopadhyay, T. & Roth, J. A. Superinduction of wild-type p53 protein after 2-methoxyestradiol treatment of Ad5p53-transduced cells induces tumor cell apoptosis. *Oncogene* **17**, 241–246 (1998).
- Fujimura, M. *et al.* Manganese superoxide dismutase mediates the early release of mitochondrial cytochrome c and subsequent DNA fragmentation after permanent focal cerebral ischemia in mice. *J. Neurosci.* **19**, 3414–3422 (1999).
- Bindokas, V. J., Jordan, J., Lee, C. C. & Miller, R. J. Superoxide production in rat hippocampal neurons: selective imaging with hydroethidine. *J. Neurosci.* **16**, 1324–1336 (1996).
- Ukeda, H., Maeda, S., Ishii, T. & Sawamura, M. Spectrophotometric assay for superoxide dismutase based on tetrazolium salt 3'-1-[(phenylamino)-carbonyl]-3,4-tetrazolium-bis(4-methoxy-6-nitro) benzenesulfonic acid hydrate reduction by xanthine-xanthine oxidase. *Anal. Biochem.* **251**, 206–209 (1997).
- Zwacka, R. M., Dudus, L., Epperly, M. W., Greenberger, J. S. & Engelhardt, J. F. Redox gene therapy protects human IB-3 lung epithelial cells against ionizing radiation-induced apoptosis. *Hum. Gene Ther.* **9**, 1381–1386 (1998).
- Heiden, M. G. V., Chandel, N. S., Williamson, E. K., Schumacker, P. T. & Thompson, C. B. Bcl-X $_L$ regulates the membrane potential and volume homeostasis of mitochondria. *Cell* **91**, 627–637 (1997).
- Marzo, I. *et al.* Bax and adenine nucleotide translocator cooperate in the mitochondrial control of apoptosis. *Science* **281**, 2027–2031 (1998).
- Li, P. *et al.* Cytochrome c and ATP-dependent formation of Apaf-1/caspase-9 complex initiates an apoptotic protease cascade. *Cell* **91**, 479–489 (1997).
- Rigo, A., Viglino, P. & Rotilio, G. Kinetic study of O $_2^{\cdot-}$ dismutation by bovine superoxide dismutase: evidence for saturation of the catalytic sites by O $_2^{\cdot-}$. *Biochem. Biophys. Res. Commun.* **63**, 1013–1018 (1975).
- D'Amato, R. J., Lin, C. M., Flynn, E., Folkman, J. & Hamel, E. 2-methoxyestradiol, an endogenous mammalian metabolite, inhibits tubulin polymerization by interacting at the colchicine site. *Proc. Natl Acad. Sci. USA* **91**, 3964–3968 (1994).
- Fotsis, T. *et al.* The endogenous oestrogen metabolite 2-methoxyestradiol inhibits angiogenesis and suppresses tumor growth. *Nature* **368**, 237–239 (1994).
- Gutteridge, J. M., Rowley, D. A. & Halliwell, B. Superoxide-dependent formation of hydroxyl radicals and lipid peroxidation in the presence of iron salts. Detection of 'catalytic' iron and anti-oxidant activity in extracellular fluids. *Biochem. J.* **206**, 605–609 (1982).
- Hamana, K. *et al.* DNA strand scission by enzymatically reduced mitomycin C: evidence for participation of the hydroxyl radical in the DNA damage. *Biochem. Int.* **10**, 301–309 (1985).
- Pervaiz, S., Ramalingam, J. K., Hirpara, J. L. & Clement, M. Superoxide anion inhibits drug-induced tumor cell death. *FEBS Lett.* **459**, 343–348 (1999).
- Casparry, W. J., Nizialek, C., Lanzetta, D. A., Friedman, R. & Bachur, N. R. Bleomycin A $_2$: a ferrous oxidase. *Mol. Pharmacol.* **16**, 256–260 (1979).
- Young, R. C., Ozols, R. F. & Myers, C. E. The anthracycline antineoplastic drugs. *N. Engl. J. Med.* **305**, 139–153 (1981).

- Gajewski, E., Rao, G., Nackerdien, Z. & Dizdaroğlu, M. Modification of DNA bases in mammalian chromatin by radiation-generated free radicals. *Biochemistry* **29**, 7876–7882 (1990).
- Mohanty, J. G., Jaffe, J. S., Schulman, E. S. & Raible, D. G. A highly sensitive fluorescent micro-assay of H $_2\text{O}_2$ release from activated human leukocytes using a dihydroxyphenoxazine derivative. *J. Immunol. Methods* **202**, 133–141 (1997).
- Huang, P., Chubb, S. & Plunkett, W. Termination of DNA synthesis by 9- β -D-arabinofuranosyl-2-fluoroadenine: a mechanism for cytotoxicity. *J. Biol. Chem.* **265**, 16617–16625 (1990).

Acknowledgements

We thank J. Engelhardt and the Vector Core Facility of University of Iowa for Ad.CuZn-SOD, Ad.MnSOD and the control viral vectors; M. Du and A. Sandoval for technical assistance in the isolation and preparation of primary cells from health donors and from leukaemia patients; and J. Richard for editorial assistance. This work was supported in part by NIH/NCI grants to P.H. and W.P.

Correspondence and requests for materials should be addressed to P.H. (e-mail: phuang@notes.mdacc.tmc.edu).

A chemical switch for inhibitor-sensitive alleles of any protein kinase

Anthony C. Bishop*, Jeffrey A. Ubersax†, Dejah T. Petsch‡, Dina P. Matheos§, Nathanael S. Gray||, Justin Blethrow†, Eiji Shimizu§, Joe Z. Tsien§, Peter G. Schultz||, Mark D. Rose§, John L. Wood‡, David O. Morgan† & Kevan M. Shokat*¶

* Department of Chemistry, Princeton University, Princeton, New Jersey 08544, USA

† Departments of Physiology and Biochemistry & Biophysics, University of California San Francisco, San Francisco, California 94143-0444, USA

‡ Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, USA

§ Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544, USA

|| Genomics Institute of the Novartis Research Foundation, San Diego, California 92121, USA

Protein kinases have proved to be largely resistant to the design of highly specific inhibitors, even with the aid of combinatorial chemistry¹. The lack of these reagents has complicated efforts to assign specific signalling roles to individual kinases. Here we describe a chemical genetic strategy for sensitizing protein kinases to cell-permeable molecules that do not inhibit wild-type kinases². From two inhibitor scaffolds, we have identified potent and selective inhibitors for sensitized kinases from five distinct subfamilies. Tyrosine and serine/threonine kinases are equally amenable to this approach. We have analysed a budding yeast strain carrying an inhibitor-sensitive form of the cyclin-dependent kinase Cdc28 (CDK1) in place of the wild-type protein. Specific inhibition of Cdc28 *in vivo* caused a pre-mitotic cell-cycle arrest that is distinct from the G1 arrest typically observed in temperature-sensitive *cdc28* mutants³. The mutation that confers inhibitor-sensitivity is easily identifiable from primary sequence alignments. Thus, this approach can be used to systematically generate conditional alleles of protein kinases, allowing for rapid functional characterization of members of this important gene family.

Cell-permeable inhibitors that are selective for individual protein kinases would allow the direct investigation of the cellular function of each kinase in a pathway. Such molecules hold several advantages over genetic strategies for the disruption of kinase activity. Unlike gene deletions, drug-like inhibitors act quickly and reversibly, and

¶ Present address: Department of Cellular and Molecular Pharmacology, University of California San Francisco, San Francisco, CA 94143-0450, USA

do not allow the cell to compensate for the missing kinase activity during development. Likewise, highly specific inhibitors are preferable to temperature-sensitive alleles, as the analysis of temperature-sensitive phenotypes may be complicated by the effects of heat shock. In addition, the mechanism of temperature-sensitive protein inactivation is rarely understood in molecular detail.

We previously engineered the ATP-binding sites of Src-family tyrosine kinases to contain unique binding pockets that are not present in any wild-type protein kinase^{2,4}. Mutation of Ile 338 to glycine in v-Src conferred unique inhibitor-sensitivity to chemically modified derivatives of PP1, the Src-family-selective inhibitor^{2,5}. We

reasoned that mutation of the equivalent position to a small amino acid may also confer unique inhibitor sensitivity in other protein kinase families. We therefore generated a series of rationally engineered protein kinases from divergent subfamilies. We selected at least one protein kinase from four distinct and physiologically important sub-families (Fig. 1a): Src family (v-Src, c-Fyn)^{6,7}, Abl family (c-Abl)⁸, Ca²⁺/calmodulin-dependent family (CAMK II α)⁹ and cyclin-dependent family (CDK2)¹⁰. The amino acid corresponding to residue 338 of v-Src was mutated to a small residue (glycine or alanine) for all of the above protein kinases to generate the following 'analogue-sensitive' (as) kinases: v-Src (I338G, v-Src-as1), c-Fyn (T339G, c-Fyn-as1), c-Abl (T315A, c-Abl-as2), CAMK II α (F89G, CAMK II α -as1) and CDK2 (F80G, CDK2-as1) (Fig. 1b; for c-Abl the T315G mutant, c-Abl-as1, was unstable; Y. Liu, personal communication). We then tested the mutation-dependent sensitivity of the analogue-sensitive kinases to two groups of derivatized inhibitors; one based on the indolocarbazole natural product (+)-K252a (Fig. 1c, 1) and one based on PP1 (Fig. 1e).

K252a is a very general protein kinase inhibitor which we predicted would bind kinase active sites in an orientation identical to that of the closely related compound, staurosporine¹¹. In the crystal structure of CDK2 bound to staurosporine, the F80 side chain is only 3.7 Å removed from C7 of staurosporine¹¹. Thus, we anticipated that developing unique indolocarbazole-based inhibitors of our sensitized kinases would require selective manipulation of C7. We synthesized a small group of C7-derivatized K252a analogues¹² (Fig. 1c, 2–6) and screened them against the wild-type and engineered protein kinases for *in vitro* inhibition of phosphorylation of an exogenous substrate (see Supplementary Information). The 2-methyl-propyl analogue (3) bound to the engineered Src-family kinases, v-Src-as1 and c-Fyn-as1, with high selectivity (≥ 3000 -fold as compared with all wild-type kinases in the group) and subnanomolar half-maximal inhibitory concentration (IC₅₀) values (230 pM and 550 pM, respectively). These values represent the most potent inhibition of any Src-family kinase reported to date. Also, the 2-phenyl-ethyl derivative (6) is a potent and moderately selective inhibitor for the mutated CDK2 and CAMK II α enzymes (see Supplementary Information). No selective inhibitors of c-Abl-as2 were identified from the K252a-based inhibitor group.

We next tested the same array of wild-type and mutant kinases against a group of C3-modified PP1 analogues² (Fig. 1e). Although PP1 is Src-family-selective, we reasoned that mutation of the 338

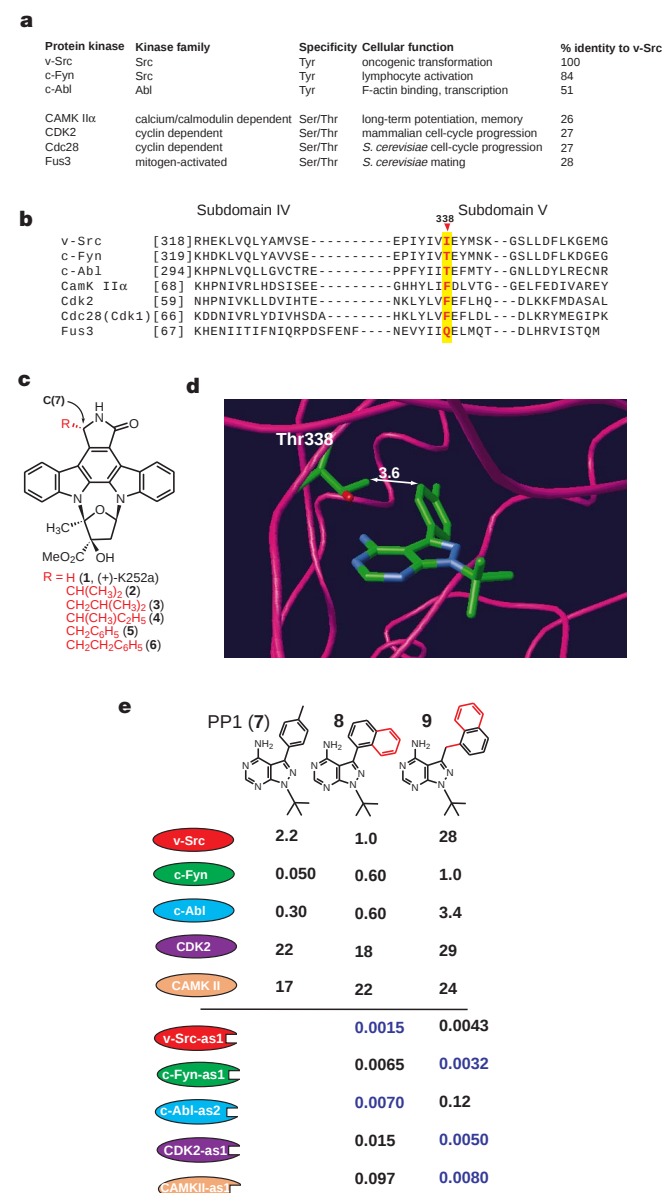


Figure 1 Structures of mutant kinase inhibitors. **a**, **b**, Substrate specificities, cellular functions, kinase domain sequence conservation and partial sequence alignment of the protein kinases. **c**, Chemical structures of (+)-K252a (1) and C7-derivatized K252a analogues (2–6). **d**, Crystal structure of PP1 bound to the Src-family tyrosine kinase Hck²⁹. The Hck peptide backbone is shown as a magenta ribbon. The atoms of the Thr 338 side chain and PP1 are coloured as follows: carbon, green; nitrogen, blue; oxygen, red. **e**, Chemical structures of PP1 (7) and C3-derivatized PP1 analogues (8, 9), and IC₅₀ values (μ M) for 7–9 against a group of wild-type and engineered protein kinases. IC₅₀ values for the best K252a derivative/engineered kinase pair for each target are shown in blue.

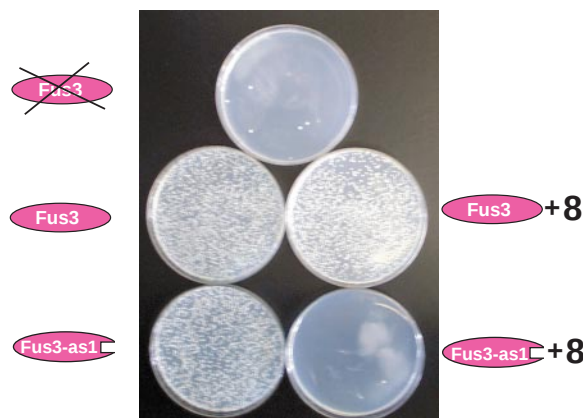


Figure 2 Selective disruption of *fus3-as1* yeast mating by analogue 8. Haploid *URA3 his3* *S. cerevisiae* expressing either wild-type Fus3, Fus3-as1, or no Fus3 were mated with an equal number of *ura3 HIS3 fus1 Δ fus2 Δ* cells and plated on media lacking uracil and histidine. Plates were inoculated with 0.1 ml of the mating cultures from the following strains (from top): *fus3 Δ* ; wild-type; wild-type + 5 μ M analogue 8; *fus3-as1*; *fus3-as1* + 5 μ M analogue 8.

position (v-Src numbering, Fig. 1d) to a small amino acid may confer unique PP1 analogue-sensitivity to other kinases, as this amino acid is partially responsible for the selectivity of PP1 itself¹³.

Analysis of the PP1 analogue inhibition data (Fig. 1e) reveals some striking trends. PP1, which is much more selective than K252a (ref. 5), yielded analogues with wider utility in the context of the engineered kinase/inhibitor pairs. Each of the five target kinases was inhibited by PP1 analogues at low nanomolar concentrations with target-specificities ranging from 85-fold to 400-fold (measured against the most inhibitable wild-type kinase). There is little or no correlation between the wild-type PP1 IC₅₀ and the PP1 analogue sensitivity of the same (engineered) kinase. This is most apparent in the Ser/Thr kinases CDK2 and CAMK II α . These wild-type enzymes both possess weak PP1 IC₅₀ values of greater than 15 μ M. However, both CDK2-as1 and CAMK II α -as1 are very potently inhibited by the C3-1'-naphthylmethyl PP1 analogue (**9**; 5.0 nM and 8.0 nM, respectively). This dichotomy is most probably due to a combination of the importance of residue 338 in determining the PP1 sensitivity of a given protein kinase¹³, and the affinity of the naphthyl ring system for the expanded kinase active site². The IC₅₀ values for analogue **9** for all of the glycine mutants were within a threefold range (all < 10 nM) of each other even though their wild-type PP1 sensitivities vary by more than 400-fold. None of the wild-type kinases tested are inhibited at concentrations of less than 1 μ M **9**, showing the high target-selectivity of this compound (Fig. 1e).

The PP1 analogues **8** and **9** are selective for different space-creating mutations on the basis of their size and flexibility. The more rigid C3-1'-naphthyl PP1 analogue (**8**) shows a broader range of potency between different kinase glycine mutants (Fig. 1e); however, it potently inhibits c-Abl-as2 (7.0 nM) with high specificity. In addition, **8** inhibits the alanine mutant of v-Src (I338A, v-Src-as2; IC₅₀ = 1.0 nM) more strongly than it inhibits the corresponding

glycine mutant, suggesting that this molecule will provide a general scheme for the development of monospecific inhibitors for protein kinases whose activity or stability is significantly compromised by the glycine mutation.

We then investigated whether our strategy could be used to generate inhibitor-sensitive kinase alleles with *in vivo* utility. PP1 analogues can be used for target-specific inhibition of retrovirally expressed v-Src in murine fibroblasts². We were interested, however, in showing monospecific inhibition of endogenous kinase gene products in an organism that has broad utility in traditional genetic

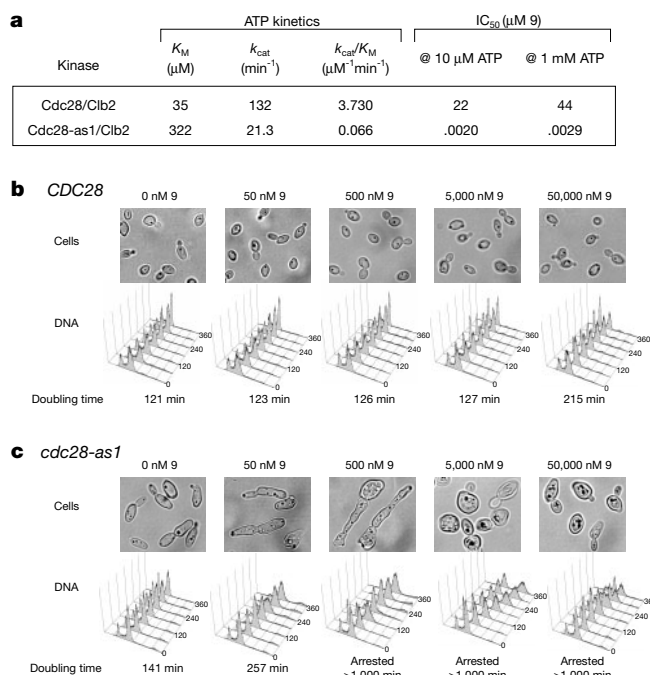


Figure 3 Characterization of the *cdc28-as1* mutant. **a**, Cdc28-as1 is highly sensitive to analogue **9** *in vitro* and is a slightly weakened kinase. **b–c**, The *cdc28-as1* strain is highly sensitive to **9** and displays several cell-cycle arrest points upon inhibitor addition. *CDC28* and *cdc28-as1* strains were grown to mid-log phase and resuspended in fresh media containing the indicated concentration of **9**. Samples were taken at the indicated times and fixed for microscopy, flow cytometric analysis and cell counting. The cell micrographs shown were taken 5 h after addition of compound **9**.

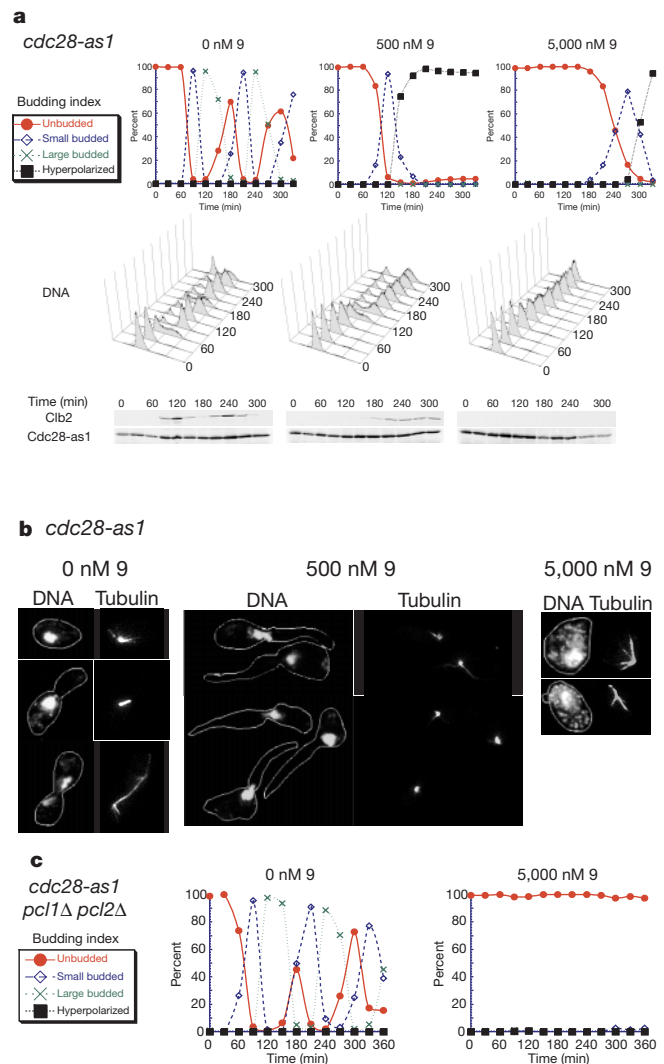


Figure 4 Characterization of *cdc28-as1* phenotype at different concentrations of analogue **9**. **a**, *cdc28-as1* cells synchronized in G1 and released into analogue **9** display a G2/M arrest (500 nM) and a G1 arrest (5,000 nM). *cdc28-as1* strains were arrested in G1 by 1 μ g ml⁻¹ α -factor for 2.5 h. Indicated concentrations of **9** were added for 20 min, and cells were released from G1 by washing and resuspension in media containing the same concentration of **9**. Every 30 min, samples were fixed for analysis of budding index and flow cytometry as in Fig. 3. Parallel samples were analysed by western blotting for Clb2 and Cdc28. **b**, *cdc28-as1* cells arrest either before mitotic entry with no mitotic spindle (500 nM) or in a G1-like state with an interphase microtubule array (5,000 nM). *cdc28-as1* cells were arrested in G1 and released as above. Cells are shown 240 min after release, except those released into 0 nM **9**, where cells were taken at 180 (top), 210 (middle) and 240 (bottom) min after release to show normal spindle morphology during the cell cycle. Tubulin and DNA staining were done as described³⁰. **c**, Budding of cells released from G1 into 5,000 nM **9** is due to Pho85/Pcl1-2 activity. G1 synchronization, release and budding indices were done as above in a *cdc28-as1* *pcl1Δ::HIS3* *pcl2Δ::URA3* strain.

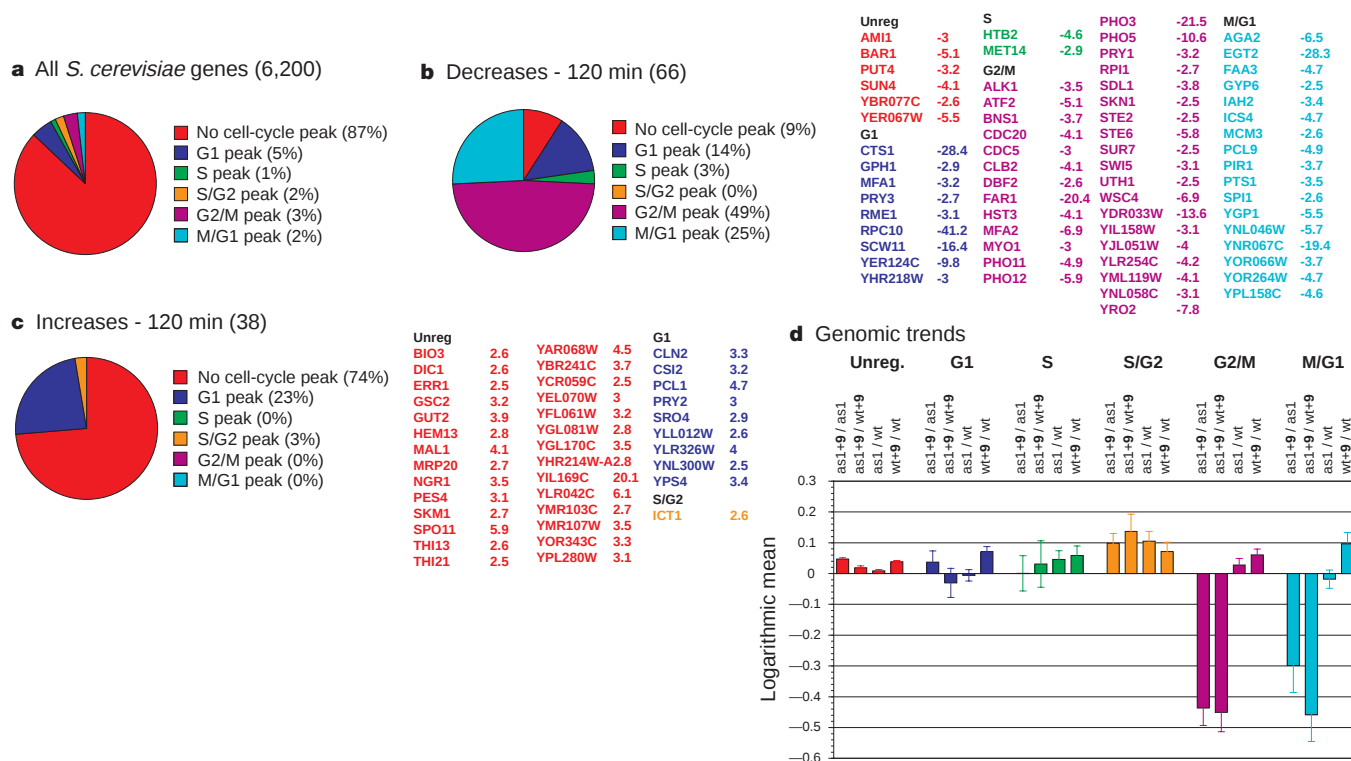
screens. The budding yeast, *Saccharomyces cerevisiae*, is widely used as a unicellular model organism in genetic studies as it is susceptible to genetic manipulation and grows rapidly. The *S. cerevisiae* genome encodes over 120 protein kinases including homologues of proteins from many mammalian kinase families¹⁴. Ideally, the kinase-sensitization strategy could be extended to these protein kinases without individually purifying and assaying every enzyme *in vitro*. To test this idea, we selected the *S. cerevisiae* MAP kinase Fus3, which is required for induction of pheromone-inducible genes, cell-cycle arrest and cell fusion during yeast mating¹⁵. No temperature-sensitive alleles of Fus3 have been isolated, possibly because of the temperature sensitivity of the mating process¹⁶.

We mutated Gln 93 of FUS3 (corresponding to I338 in v-Src) to glycine (Fus3-as1), and expressed this enzyme in budding yeast that lack wild-type Fus3. The *fus3-as1* mutant fully complemented the gene deletion, as shown by the mating of equal numbers of haploid wild-type or *fus3-as1* cells (*URA3 his3*) to a *fus1Δ fus2Δ ura3 HIS3* strain followed by selection for diploid progeny on media lacking uracil and histidine. Wild-type and *fus3-as1* strains both yielded about 7.5×10^4 colony-forming units (c.f.u.) per ml while mating of a *fus3Δ* strain gave only 0–100 c.f.u. per ml under the same selection conditions. When the mating of a *fus3-as1* strain was carried out in the presence of 50 μ M or 5 μ M **8**, the number of diploids formed was indistinguishable from a *fus3Δ* strain (Fig. 2). Even at 500 nM **8**, the *fus3-as1* strain gave only 1.7×10^3 c.f.u. per ml, a 44-fold decrease from equivalent control cells. Because of the competition with millimolar quantities of ATP in the yeast cell, this strong inhibition at 500 nM analogue **8** implies an *in vitro* IC₅₀ for Fus3-as1 in the low nanomolar range². Analogue **9** also disrupted mating in

fus3-as1 cells, but with less potency (see Supplementary Information). By contrast, no decrease in mating efficiency was observed when wild-type cells were treated with **8** or **9** ($(0.6-1.1) \times 10^5$ c.f.u. per ml at all concentrations up to 50 μ M). Therefore *fus3-as1* represents the first conditional allele of this MAP kinase. Through the *in vivo* addition of analogue **8**, the activity of the Fus3 kinase can be controlled selectively, rapidly and in a dose-dependent fashion.

We next set out to use PP1 analogues to study the *S. cerevisiae* cell cycle through selective inhibition of the protein kinase Cdc28. This enzyme is the principal CDK in budding yeast and is essential for progression through several stages of the cell cycle¹⁷. The function of Cdc28 has been studied primarily by the analysis of temperature-sensitive mutant alleles³. At 37 °C, most of these mutants arrest in G1, suggesting that progression through START (the G1 commitment point) is sensitive to inhibition of Cdc28 activity. Notably, analysis of temperature-sensitive mutants has not provided clear evidence that Cdc28 is involved in the G2/M transition, despite abundant evidence that CDK1 is required for mitotic entry in other eukaryotes¹⁸. One well-studied temperature-sensitive mutant of *CDC28*, *cdc28-1N*, exhibits a metaphase arrest at the restrictive temperature¹⁹; however, both structural and biochemical data have suggested that the arrest of *cdc28-1N* cells is due to a temperature-sensitive protein–protein interaction (with Cks1), and not a general block of Cdc28 kinase activity^{19,20}.

Cdc28 is 62% identical to human CDK2, suggesting that the engineered Cdc28, Cdc28-as1 (F88G) should be susceptible to inhibition by C3-derivatized PP1 analogues. We expressed and purified Cdc28–His₆ and Cdc28-as1–His₆ from Sf9 insect cells, and formed active complexes with bacterially expressed MBP–Clib2.



Relative to wild-type Cdc28, Cdc28-as1 displayed a moderate reduction in activity, including a tenfold reduction in ATP-binding affinity and a sixfold reduction in maximum ATP turnover rate (Fig. 3a). More notably, Cdc28-as1 was highly sensitive to **9**. In 1 mM ATP, which approximates intracellular ATP concentration, Cdc28-as1 was about 15,000-fold more sensitive to **9** than was wild-type Cdc28 ($IC_{50} \approx 3$ nM for Cdc28-as1, and 44,000 nM for Cdc28). Thus, the single substitution of a glycine for a phenylalanine results in a Cdc28 mutant that displays both high affinity and specificity for analogue **9**.

To examine the function of Cdc28 *in vivo*, we constructed a yeast strain in which the wild-type copy of *CDC28* was replaced by *cdc28-as1*. In the absence of **9**, *cdc28-as1* cells displayed normal viability and growth on plates, although they were hyperpolarized and larger than an isogenic *CDC28* strain, and had a 20% longer doubling time in liquid culture (Fig. 3b, c). Flow cytometric analysis of asynchronously dividing cells revealed that the *cdc28-as1* and *CDC28* strains displayed similar DNA content profiles (Fig. 3b, c). We also used synthetic oligonucleotide DNA arrays to measure genome-wide transcriptional differences between asynchronous wild-type and *cdc28-as1* cells²¹. In two separate experiments, greater than twofold changes were observed in only 11 transcripts (0.2% of the genome).

We suspect that the minor growth defect in *cdc28-as1* cells is due to the sixfold lower catalytic rate constant (k_{cat}) of the mutant enzyme, which would result in a sixfold lower activity at the high ATP concentrations *in vivo* (the lower ATP-binding affinity of the mutant should not be relevant *in vivo*, where ATP concentration is much greater than the Michaelis constant, K_m). Thus, *cdc28-as1* is a slightly weakened allele of *CDC28*, but can support cell-cycle progression.

We analysed the effects of analogue **9** on cell growth and morphology. Proliferation of wild-type cells was unaffected by the inhibitor except at very high concentrations (50,000 nM), when doubling times increased about twofold (Fig. 3b). Hybridization to oligonucleotide arrays revealed that transcription of only six genes (0.1%) was changed more than twofold after 30 min treatment of wild-type cells with 500 nM compound **9**. Treatment of cells for 120 min caused even fewer transcriptional responses: only three transcripts showed more than twofold changes. No significant activation of stress-responsive transcripts was observed. We conclude that treatment with 500 nM compound **9** has no significant effects on wild-type cell physiology.

The single amino-acid change in the Cdc28-as1 kinase makes it very sensitive to **9** *in vivo*. Growth of the *cdc28-as1* strain was inhibited 50% at 50–100 nM **9**; complete growth arrest occurred at 500 nM **9** (Fig. 3c). The close correspondence between the IC_{50} *in vivo* and those measured in kinase assays (Fig. 3a) illustrates the efficacy with which **9** can penetrate a yeast cell, a feature not shown by other small-molecule CDK inhibitors¹.

Analysis of DNA content and morphology of *cdc28-as1* cells showed that lower concentrations of **9** caused a delay (50 nM) or arrest (500 nM) with a G2/M DNA content and large hyperpolarized buds (Fig. 3c). Higher inhibitor concentrations (5,000 nM) caused a non-uniform arrest that included unbudded G1 cells as well as large-budded G2/M cells; bud hyperpolarization was no longer evident. The *cdc28-as1* mutant phenotype is recessive: the inhibitor had no effect in diploid cells carrying both a wild-type *CDC28* gene and a *cdc28-as1* mutant allele (data not shown). Thus, the effects of **9** in *cdc28-as1* cells are not due to dominant-negative effects of inactive kinase complexes in the cell.

When *cdc28-as1* cells were synchronized in G1 with the yeast mating pheromone, α -factor, and released into fresh media containing 500 nM **9**, initiation of budding and DNA synthesis was delayed about 30–60 min, and accumulation of the mitotic cyclin Clb2 was delayed 60–90 min (Fig. 4a). Cells went on to arrest with moderate Clb2 levels, highly hyperpolarized buds and a G2/M DNA content, much like the arrest observed in asynchronous cells.

Microscopic analysis of microtubules and DNA revealed that these cells lacked a mitotic spindle and arrested with a single DNA mass correctly positioned at the bud neck (Fig. 4b). Cells treated with low concentrations of the inhibitor are therefore slightly delayed during progression through early stages of the cell cycle, and eventually arrest because of a failure to assemble a mitotic spindle and enter mitosis.

When we released G1-synchronized *cdc28-as1* cells into media containing 5,000 nM **9**, they arrested with a 1C DNA content (unreplicated DNA) and initially failed to bud, suggesting an arrest at START (this arrest was reversible, as washing away the inhibitor after 60 min allowed cells to rapidly resume a normal cell cycle; see Supplementary Information). Arrested cells eventually budded after 180–240 min of inhibitor treatment, and arrested with a 1C DNA content, large hyperpolarized buds, a single DNA mass in the mother cell, an interphase astral microtubule array and no detectable Clb2 (Fig. 4a, b).

Pho85, a CDK that is closely related to Cdc28, has been implicated in the initiation of budding in *S. cerevisiae*²². We therefore investigated whether the budding that we observed in *cdc28-as1* cells released into 5,000 nM **9** was due to residual Cdc28-as1 or Pho85 activity. We constructed a *cdc28-as1* strain lacking *PCL1* and *PCL2*, which encode the two G1 activating cyclins of Pho85. When these cells were synchronized in G1 and released into 5,000 nM **9**, they arrested with a 1C DNA content and did not bud, even after 6 h (Fig. 4c). We conclude that Pcl1- and Pcl2-associated Pho85 activities, and not residual Cdc28 activity, are responsible for the greatly delayed budding observed in *cdc28-as1* cells treated with high concentrations of inhibitor.

To characterize further the cell-cycle arrest caused by inhibitor treatment, we analysed genome-wide transcription after treatment of asynchronous *cdc28-as1* cells with 500 nM **9** (Fig. 5). Treatment for 2 h caused 2.5-fold or greater changes in the transcription of 104 genes (66 decreases, 38 increases) (Fig. 5b, c). Of the downregulated transcripts, 60 are cell-cycle-regulated and 32 have peak expression at G2/M²³. This downregulated G2/M subset contains a broad range of well-established mitotic regulators, including *CLB2*, *SWI5*, *CDC20* and *CDC5*. Decreased levels of the *CLB2* transcript are consistent with the delay in Clb2 accumulation after release from G1 arrest (Fig. 4a). Treatment with **9** for 30 min produced a similar G2/M transcriptional block (37 of the 42 downregulated genes are cell-cycle regulated, and 57% of these peak at G2/M; see Supplementary Information), suggesting that the inhibitor reaches active intracellular concentrations within minutes after addition to the growth media.

Two-hour drug exposure also increased the expression of 38 transcripts over 2.5-fold (10 cell-cycle regulated) (Fig. 5c). Nine of the ten cell-cycle regulated transcripts are expressed at peak levels in G1. This group includes genes encoding G1 cyclins (Cln2, Pcl1). This result, combined with the fact that the Cln–Cdc28 inhibitor Far1 was downregulated 20-fold, suggests that prolonged Cdc28 inhibition and G2/M arrest lead to a transcriptional program that boosts the activity of G1 cyclin/CDK complexes.

To assess general trends in the transcriptional response to Cdc28 inhibition, we also calculated mean changes in the expression of all genes from each of the principal cell-cycle gene clusters (Fig. 5d). This analysis confirmed that inhibitor treatment of *cdc28-as1* cells led primarily to decreased expression of genes that are normally expressed in the G2/M and M/G1 stages of the cell cycle.

We conclude that low concentrations of **9** cause *cdc28-as1* cells to arrest before mitosis with a phenotype similar to that observed in cells lacking the mitotic cyclins Clb1–Clb4 (ref. 24). Entry into mitosis appears particularly sensitive to inhibition of Cdc28 catalytic activity, and only at higher concentrations of inhibitor does complete inhibition of passage through G1 occur. These data are generally consistent with the idea that different cell-cycle events are triggered by specific threshold levels of CDK activity, and that later

events require higher amounts of activity²⁵.

Our results do not seem consistent with previous studies^{3,17} of temperature-sensitive *cdc28* mutants, most of which arrest as unbudded cells in G1. One explanation for this discrepancy is that there are major differences in the inhibitor sensitivity of different Cdc28/cyclin complexes. We have been unable to test this possibility directly because we find that Cdc28-as1/Cln complexes exhibit insufficient activity *in vitro* to allow detailed analysis of their inhibitor-sensitivity (data not shown). Nevertheless, it seems unlikely that Cdc28-as1/Cln complexes exhibit significantly reduced sensitivity to **9**, given that four different as1 kinase mutants (from three different kinase families) displayed roughly the same nanomolar sensitivity that we observed for Cdc28-as1/Clb2 (Figs 1 and 3a).

Using a chemical switch to design a unique protein/small molecule interface we can systematically engineer protein kinases with unique sensitivity to cell-permeable inhibitors. This approach can be successful even for protein kinases that are not potently inhibited by the chosen 'lead' inhibitor, and different small-molecule scaffolds can be used to generate analogues that are specific for the same target. Developing several chemical structures for inhibition of a specific target may prove valuable for application of this kinase-sensitization strategy to mammalian studies, as bio-availability and selectivity requirements (>1000 protein kinases) become more stringent. Our data suggest that most protein kinases will be susceptible to this target-specific inhibition strategy. For organisms that readily undergo homologous recombination, this approach raises the possibility of generating conditional allelic strains for every protein kinase in the genome. A library of such strains would represent a significant step forward in realizing the pharmacogenomic vision of identifying small-molecule ligands for every gene product in the cell. □

Methods

Chemical synthesis

(+)-K252a (ref. 12) and PP1 (ref. 2) analogues were synthesized as described.

Protein expression and purification

Wild-type and engineered GST-v-Src (catalytic domain), GST-c-Fyn, GST-c-Abl, GST-CAMK II α and CDK2-His₆ were expressed and purified as described¹³.

Lysate was prepared from insect cells co-infected with baculoviruses encoding Cdc28-His₆ or Cdc28-as1-His₆ and Cak1-HA₃ (ref. 26). Cdc28-His₆ and Cdc28-as1-His₆ were purified by metal-affinity, followed by anion-exchange (Pharmacia SP Sepharose Fast Flow) and cation-exchange (Pharmacia Q Sepharose Fast Flow) chromatography. MBP-Clb2 was purified from lysates of bacteria expressing MBP-Clb2 (a gift from R. Deshaies).

Yeast plasmid and strain construction

Genetic and microbial techniques were done essentially as described²⁷. To generate a plasmid containing *FUS3*, we constructed plasmid pMR3980 by cloning an *EcoRI*/PmlI fragment from pMR2412 into pRS416 (ref. 28) cut with *EcoRI*/SmaI. The Q93G mutant allele, pMR4280, was created by phagemid *in vitro* mutagenesis of plasmid pMR3980. Strain MY2914, MAT α fus3 Δ :LEU2, was transformed with either pMR3980, pMR4280 or pRS416.

All *CDC28* strains were derivatives of W303 (*ura3-1*, *leu2-3,112*, *trp1-1*, *his3-11,15*, *ade2-1*, *can1-100*, *GAL+*). To create *cdc28-as1*, the *CDC28* coding sequence and 400 bp of 5'- and 386 bp of 3' flanking DNA were PCR-amplified from genomic DNA and ligated into pRS306 (ref. 28) to make pJAU1. The F88G mutation was engineered by oligonucleotide-directed mutagenesis of pJAU1. The plasmid was integrated into the wild-type *CDC28* locus following *AflIII* digestion by a pop-in-pop-out strategy as described²⁷.

In vitro kinase assays

We carried out *in vitro* kinase assays (except for Cdc28) in the presence of 0.2 μ Ci μ l⁻¹ [γ -³²P]ATP at low ATP concentration (10 nM) so that IC₅₀ values represent a rough measure¹³ of the inhibition constant (K_i). Measurement of inhibitor IC₅₀ values were done as described¹³.

Purified Cdc28-His₆ (1 nM) and MBP-Clb2 (3 nM) were incubated for 10 min at 23 °C in a 25 μ l reaction mixture containing 5 μ g histone H1, 1 μ Ci of [γ -³²P]ATP (1 μ Ci per 10 μ M and 1 μ Ci per 1 mM), and varying concentrations of compound **9** in kinase buffer (25 mM HEPES-NaOH pH 7.4, 10 mM NaCl, 10 mM MgCl₂ and 1 mM dithiothreitol). Reaction products were analysed by 15% SDS-PAGE followed by autoradiography. For the determination of Cdc28 kinetic constants, varying concentrations of [γ -³²P]ATP (1 μ Ci per 100 μ M) were incubated and analysed as above.

Fus3-dependent mating assays

Haploid *URA3 his3 S. cerevisiae* expressing either wild-type Fus3, Fus3-as1, or no Fus3 at an absorbance of 0.5 at 600 nm were mated with an equal number of *ura3 HIS3 fus1 Δ fus2 Δ* cells and pipetted onto a nitrocellulose disk. The disk was placed on a YPD plate containing varying amounts amount of inhibitor for 5 h at 30 °C. Cells were liberated and dilutions of the resulting cultures were plated on media lacking uracil and histidine and grown at 30 °C and the colonies were counted. All cultures were also plated on YPD media. No significant decrease in c.f.u. per ml on YPD plates was observed for any of the three strains in the presence of analogue **8** or **9**.

DNA flow cytometry

Roughly 10⁷ cells for each sample were fixed in 70% ethanol, resuspended in 50 mM Tris-HCl pH 8.0, briefly sonicated, digested with 2 mg ml⁻¹ RNase for 2 h at 37 °C, and resuspended in 0.2 ml protease solution (5 mg ml⁻¹ pepsin, 0.5% concentrated HCl) and digested for 45 min at 37 °C. We stained DNA with 1 μ M Sytox Green (Molecular Probes) in 50 mM Tris-HCl pH 7.5, and scanned 20,000 cells from each sample with a FACScan machine (Becton-Dickinson).

Genome-wide transcriptional analysis

We measured genome-wide transcriptional differences by oligonucleotide microarray analysis of cellular messenger RNA as described²¹. Changes were deemed significant if they were greater than or equal to 2.0-fold or if the transcript changed present/absent status (as indicated by Affymetrix software) in the following comparisons: for nonspecific drug effects, *CDC28* + **9** versus *CDC28*; for *cdc28-as1* effects, *cdc28-as1* versus *CDC28*. For specific Cdc28 inhibition effects, changes were deemed significant if they were equal to or greater than 2.5-fold for *cdc28-as1* + **9** versus *cdc28-as1* and also greater than or equal to 2.0-fold for *cdc28-as1* + **9** versus *CDC28* + **9**.

Greater than twofold differences were observed in expression of the following genes in a comparison of wild-type and *cdc28-as1* cells (11 overlaps between the 2 experiments are underlined). Experiment 1: *LEU2*, *YDL241W*, *YFL057C*, *YHR071W*, *CBP1*, *YJR130C*, *CWP1*, *YLL060C*, *ATRI*, *YML128C*, *YMR095C*, *YMR096W*, *YNL275W*, *YNR065C*, *ARG1*, *YOL101C*, *SPS4*, *SSU1*, *SVS1*, *OYE3*, *RLF2* and *YPR203W*. Experiment 2: *DURI*, *CIT2*, *YDL241W*, *INO2*, *YFL056C*, *YFL057C*, *YGL088W*, *YHR071W*, *TWT1*, *YHR209W*, *YIL011W*, *YIL023C*, *DAL4*, *YKL218C*, *YLL060C*, *YLR108C*, *YLR237W*, *YLR437C*, *YML071C*, *ATRI*, *YMR095C*, *YMR096W*, *YNR065C*, *ARG1*, *YOR302W*, *SPS4*, *YPL088W*, *SVS1*, *YPR013C* and *CTR1*.

Of the six transcripts that changed in wild-type cells after 30 min of 500 nM inhibitor treatment, three have no known function (YGR035C, YLR346C, YPL222W) and the others have roles in heat shock (SSA4), osmotic stress response (*GRE2*) and drug resistance (*YOR1*). The transcription of only one of these genes (YGR035C) is cell-cycle regulated²³. Treatment of wild-type cells with 500 nM inhibitor for 120 min resulted in changes in only three genes (YGR035C and two genes encoding ribosomal proteins: *RPS26A* and *RPL26B*).

Received 26 April; accepted 24 July 2000.

- Gray, N. S. *et al.* Exploiting chemical libraries, structure, and genomics in the search for kinase inhibitors. *Science* **281**, 533–538 (1998).
- Bishop, A. C. *et al.* Generation of monospecific nanomolar tyrosine kinase inhibitors via a chemical genetic approach. *J. Am. Chem. Soc.* **121**, 627–631 (1999).
- Reed, S. I. The selection of *S. cerevisiae* mutants defective in the start event of cell division. *Genetics* **95**, 561–577 (1980).
- Liu, Y., Shah, K., Yang, F., Witucki, L. & Shokat, K. M. Engineering Src family protein kinases with unnatural nucleotide specificity. *Chem. Biol.* **5**, 91–101 (1998).
- Hanke, J. H. *et al.* Discovery of a novel, potent, and Src family-selective tyrosine kinase inhibitor. *J. Biol. Chem.* **271**, 695–701 (1996).
- Brown, M. T. & Cooper, J. A. Regulation, substrates and functions of src. *Biochim. Biophys. Acta* **1287**, 121–149 (1996).
- Resh, M. D. Fyn, a Src family tyrosine kinase. *Int. J. Biochem. Cell Biol.* **30**, 1159–1162 (1998).
- Laneville, P. Abl tyrosine protein kinase. *Semin. Immunol.* **7**, 255–266 (1995).
- Kelly, P. T. Calmodulin-dependent protein kinase II. Multifunctional roles in neuronal differentiation and synaptic plasticity. *Mol. Neurobiol.* **5**, 153–177 (1991).
- Morgan, D. O. Principles of CDK regulation. *Nature* **374**, 131–134 (1995).
- Lawrie, A. M. *et al.* Protein kinase inhibition by staurosporine revealed in details of the molecular interaction with CDK2. *Nature Struct. Biol.* **4**, 796–801 (1997).
- Wood, J. L., Stoltz, B. M., Dietrich, H. J., Pflum, D. A. & Petsch, D. T. Design and implementation of an efficient synthetic approach to furanosylated indolocarbazoles: total synthesis of (+)- and (–)-K252a. *J. Am. Chem. Soc.* **119**, 9641–9651 (1997).
- Liu, Y. *et al.* Structural basis for selective inhibition of Src family kinases by PP1. *Chem. Biol.* **6**, 671–678 (1999).
- Hunter, T. & Plowman, G. D. The protein kinases of budding yeast: six score and more. *Trends Biochem. Sci.* **22**, 18–22 (1997).
- Fujimura, H. A. Yeast homolog of mammalian mitogen-activated protein kinase, Fus3/Dac2 kinase, is required both for cell fusion and for G1 arrest of the cell cycle and morphological changes by the *cdc37* mutation. *J. Cell Sci.* **107**, 2617–2622 (1994).
- Doi, S. & Yoshimura, M. Temperature-sensitive loss of sexual agglutinability in *Saccharomyces cerevisiae*. *Arch. Microbiol.* **114**, 287–288 (1977).
- Mendenhall, M. D. & Hodge, A. E. Regulation of Cdc28 cyclin-dependent protein kinase activity during the cell cycle of the yeast *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **62**, 1191–1243 (1998).
- King, R. W., Jackson, P. K. & Kirschner, M. W. Mitosis in transition. *Cell* **79**, 563–571 (1994).
- Surana, U. *et al.* The role of CDC28 and cyclins during mitosis in the budding yeast *S. cerevisiae*. *Cell* **65**, 145–161 (1991).

20. Bourne, Y. *et al.* Crystal structure and mutational analysis of the human CDK2 kinase complex with cell cycle-regulatory protein CksHs1. *Cell* **84**, 863–874 (1996).
21. Wodicka, L., Dong, H., Mittmann, M., Ho, M. H. & Lockhart, D. J. Genome-wide expression monitoring in *Saccharomyces cerevisiae*. *Nature Biotechnol.* **15**, 1359–1367 (1997).
22. Espinoza, F. H., Ogas, J., Herskowitz, I. & Morgan, D. O. Cell cycle control by a complex of the cyclin HCS26 (PCL1) and the kinase PHO85. *Science* **266**, 1388–1391 (1994).
23. Spellman, P. T. *et al.* Comprehensive identification of cell cycle-regulated genes of the yeast *Saccharomyces cerevisiae* by microarray hybridization. *Mol. Biol. Cell* **9**, 3273–3297 (1998).
24. Fitch, I. *et al.* Characterization of four B-type cyclin genes of the budding yeast *Saccharomyces cerevisiae*. *Mol. Biol. Cell* **3**, 805–818 (1992).
25. Stern, B. & Nurse, P. A quantitative model for the cdc2 control of S phase and mitosis in fission yeast. *Trends Genet.* **12**, 345–350 (1996).
26. Farrell, A. & Morgan, D. O. Cdc37 promotes the stability of protein kinases Cdc28 and Cak1. *Mol. Cell Biol.* **20**, 749–754 (2000).
27. Sherman, F., Fink, G. & Lawrence, C. *Methods in Yeast Genetics* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1974).
28. Sikorski, R. S. & Hieter, P. A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*. *Genetics* **122**, 19–27 (1989).
29. Schindler, T. *et al.* Crystal structure of Hck in complex with a Src family-selective tyrosine kinase inhibitor. *Mol. Cell* **3**, 639–648 (1999).
30. Biggins, S. *et al.* The conserved protein kinase Ipl1 regulates microtubule binding to kinetochores in budding yeast. *Genes Dev.* **13**, 532–544 (1999).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank B. Schulman and E. Harlow for early characterization of mutant cyclin-dependent kinases; C. Co, V. Nguyen and J. Li for help with flow cytometry; D. Toczyski for help with the Coulter counter; A. Su for help with transcript array analysis; S. Jaspersen, C. Carroll, C. Takizawa, E. Weiss, S. Biggins, A. Szidon, A. Rudner, D. Kahne and A. Murray for helpful advice and reagents; E. O'Shea and R. Deshaies for strains and reagents. This work was supported by funding from the National Institute of Allergy and Immunology (K.M.S.), GlaxoWellcome (K.M.S. and J.L.W.), National Institute of General Medical Sciences (D.O.M.), National Institutes of Health (J.L.W.), Bristol-Myers Squibb (J.L.W.), Yamanouchi (J.L.W.) and a predoctoral fellowship from the National Science Foundation (J.A.U.). K.M.S. is a Pew, Searle, and a Sloan Foundation Scholar.

Correspondence and requests for material should be addressed to K.M.S. (e-mail: shokat@cmp.ucsf.edu).

The protein Aly links pre-messenger-RNA splicing to nuclear export in metazoans

Zhaolan Zhou^{*,†}, Ming-juan Luo^{*,†}, Katja Straesser[‡], Jun Katahira[‡], Ed Hurt[‡] & Robin Reed^{*}

^{*} Department of Cell Biology, Harvard Medical School, Boston, Massachusetts 02115, USA

[‡] University of Heidelberg, Biochemie-Zentrum Heidelberg (BZH), Im Neuenheimer Feld 328, D-69120 Heidelberg, Germany

In metazoans, most pre-messenger RNAs contain introns that are removed by splicing. The spliced mRNAs are then exported to the cytoplasm. Recent studies showed that splicing promotes efficient mRNA export¹, but the mechanism for coupling these two processes is not known. Here we show that Aly, the metazoan homologue of the yeast mRNA export factor Yra1p (ref. 2), is recruited to messenger ribonucleoprotein (mRNP) complexes generated by splicing. In contrast, Aly does not associate with mRNPs assembled on identical mRNAs that already have no introns or with heterogenous nuclear RNP (hnRNP) complexes. Aly is recruited during spliceosome assembly, and then becomes tightly associated with the spliced mRNP. Aly shuttles between the nucleus and cytoplasm, and excess recombinant Aly increases

both the rate and efficiency of mRNA export *in vivo*. Consistent with its splicing-dependent recruitment, Aly co-localizes with splicing factors in the nucleus. We conclude that splicing is required for efficient mRNA export as a result of coupling between the splicing and the mRNA export machineries.

Pre-mRNA splicing generates an mRNP complex distinct from that assembled on the identical mRNA lacking introns (Δ i-mRNA)^{1,3}. When these two types of complexes are injected into *Xenopus* oocyte nuclei, the mRNA in the spliced mRNP is exported to the cytoplasm more efficiently than that in the Δ i-mRNP¹. This

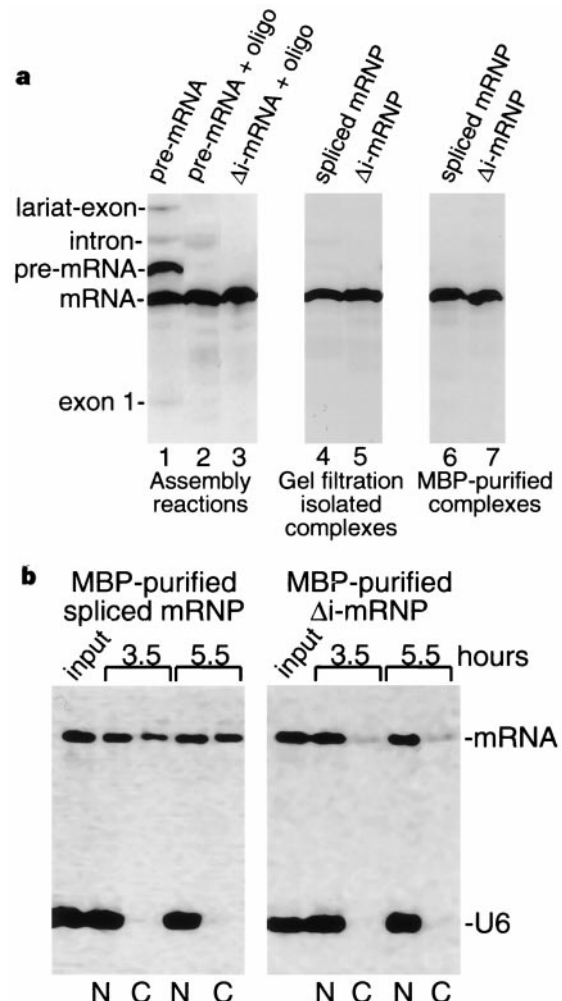


Figure 1 Purification of spliced messenger ribonucleoprotein (mRNP) that is functional in messenger RNA export. **a**, Maltose-binding protein (MBP)-affinity purification of spliced mRNP or Δ i-mRNP. AdML pre-mRNA was incubated in nuclear extracts under splicing conditions for 60 minutes (lane 1). AdML pre-mRNA or Δ i-mRNA was incubated under splicing conditions for 60 minutes, followed by addition, to both reactions, of a 12-mer oligonucleotide (5'-CTGCCTTAGTG-3') complementary to intron sequences immediately downstream from the 5' splice site (lanes 2 and 3). Incubation was continued for 30 minutes, resulting in RNase H digestion of the pre-mRNA, intron lariat (a branched intermediate structure in the intron), and lariat-intermediate (compare lanes 1 and 2). There was no effect on the Δ i-mRNA (lane 3). The complexes shown in lanes 2 and 3 were then isolated by gel filtration on Sephacryl S-500 (ref. 30). Total RNA isolated from the pooled fractions containing the mRNPs is shown in lanes 4 and 5. Gel filtration-isolated mRNPs were further purified by MBP-affinity chromatography. The total RNA in the elutions is shown in lanes 6 and 7. In negative control purification, no RNA or protein was detected in the elution, indicating the high specificity of the MBP-affinity purification strategy (data not shown). **b**, Export assays of the MBP-affinity-purified mRNPs. The spliced mRNP or Δ i-mRNP (purified as in **a**) was mixed with U6 snRNA and injected into *Xenopus* oocyte nuclei. Oocytes were incubated at 18 °C for 0 (input), 3.5 or 5.5 hours, and the total RNA from the nucleus (N) or cytoplasm (C) was analysed.

[†] These authors contributed equally to this work.

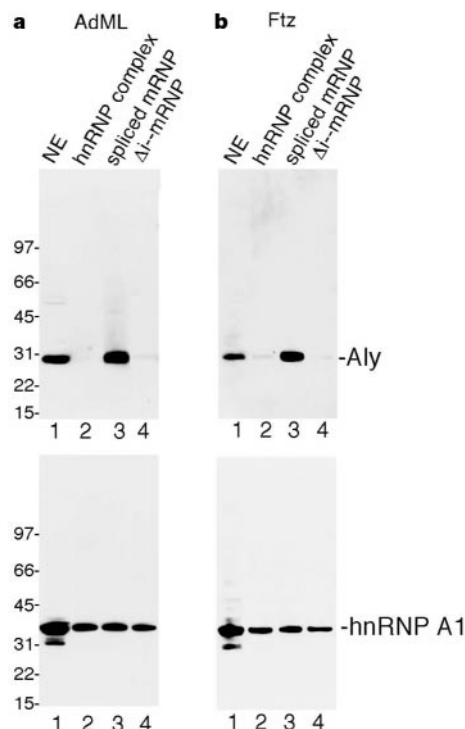


Figure 2 Splicing-dependent recruitment of Aly to the spliced mRNPs. Western blots of MBP-affinity-purified mRNPs or the hnRNP complex H, each containing 10 ng of spliced mRNA, Δ i-mRNA or pre-mRNA, were probed with the indicated antibodies. An aliquot of nuclear extract (NE, 1.0 μ l) was loaded in lane 1. Complexes were assembled on AdML (a) or Ftz RNA (b). The positions of M_r markers (K) are shown at the left.

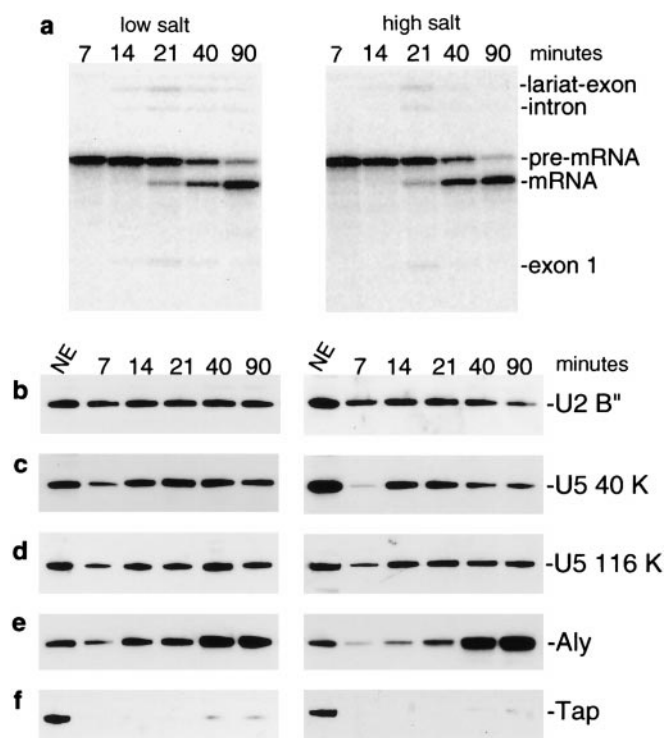


Figure 3 Timing, specificity and stability of Aly recruitment. AdML pre-mRNA was incubated under splicing conditions for the indicated times. The complexes at each time point were isolated by gel filtration followed by MBP-affinity purification in low (60 mM) or high (250 mM) salt conditions. a, RNA from the eluted complexes was analysed. b–f, Western blots of total protein from the eluted complexes shown in a were probed with the indicated antibodies.

increased efficiency could be due to the recruitment of export factors to the spliced mRNP. To test this possibility, we purified the mRNPs. The spliced mRNP was first enriched in the splicing reaction, using oligonucleotide-directed RNase H digestion to eliminate contaminating pre-mRNA (Fig. 1a, compare lanes 1 and 2). The spliced mRNP or Δ i-mRNP was then separated by gel filtration followed by maltose-binding protein (MBP)-affinity purification⁴ (Fig. 1a, lanes 4–7).

To determine whether the purified spliced mRNP and Δ i-mRNP are distinct in an mRNA export assay, the complexes were injected

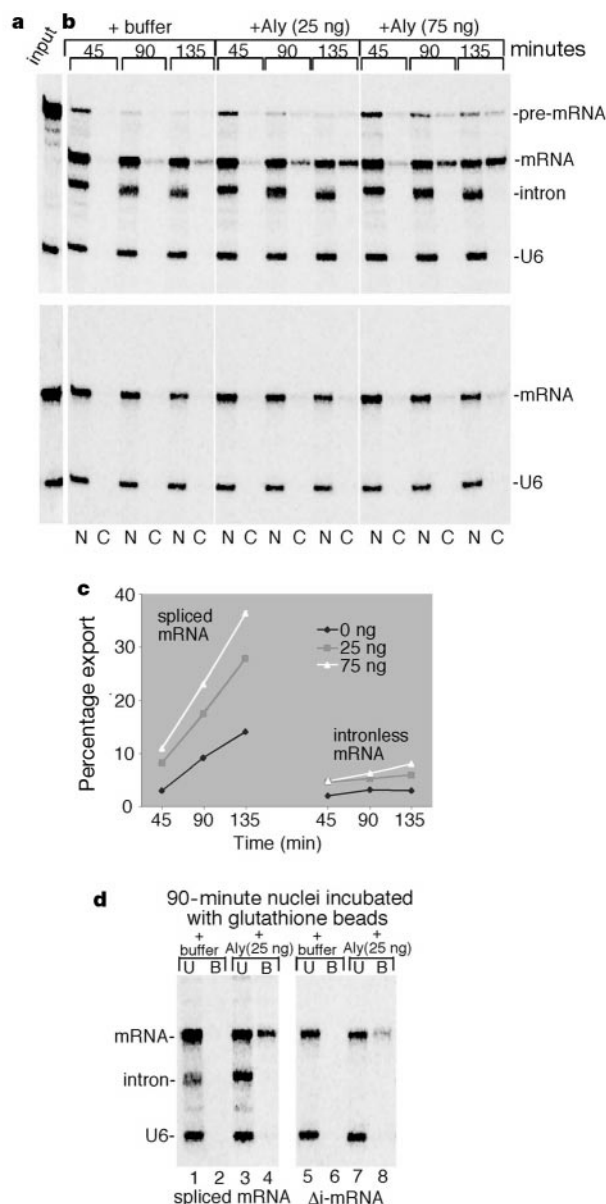


Figure 4 Aly promotes mRNA export. a, AdML pre-mRNA or Δ i-mRNA was incubated under splicing conditions *in vitro* for 20 minutes to assemble the spliceosome or Δ i-mRNP, respectively. Aliquots of the reactions were mixed with U6 snRNA, and the total RNA was analysed. b, GST–Aly (25 or 75 ng) or buffer alone was injected into the *Xenopus* oocyte cytoplasm followed by overnight incubation to import Aly. Complexes in a were injected into the nuclei of these oocytes, followed by incubation for 45, 90 and 135 minutes. RNA from the nucleus (N) or cytoplasm (C) was analysed. c, The percentage export was calculated from the data in b. d, The nuclei from the 90-minute time point in b (+ buffer and +Aly (25 ng)) were isolated, homogenized in binding buffer (20 mM HEPES–NaOH, pH 7.9, 60 mM KCl, 0.1% Triton X-100) and rotated with glutathione beads overnight. The total RNA in bound and unbound fractions was analysed.

into *Xenopus* oocyte nuclei. U6 small nuclear RNA (snRNA), which is retained in the nucleus⁵, was added to the injection mixture as a marker for intact nuclei (Fig. 1b, lanes input). As shown in Fig. 1b, the mRNA in the spliced mRNP is more rapidly and efficiently exported than that in the Δ i-mRNP. Thus, if the spliced mRNP contains a factor that promotes mRNA export, this factor is associated with the spliced mRNP after the multi-step purification.

We next investigated whether the purified spliced mRNP is associated with an mRNA export factor. Although no metazoan factors have been shown to be required for cellular mRNA export, mammalian homologues of essential *Saccharomyces cerevisiae* mRNA export proteins have been identified. In particular, Tap and Aly (also known as mREF1-I, ref. 6) are homologues of the *S. cerevisiae* export factors Mex67p and Yra1p, respectively^{2,6,7}. Mex67p interacts directly with Yra1p, and their metazoan counterparts also interact^{2,6}. Moreover, Tap binds to the constitutive transport element (CTE) which promotes export of unspliced pre-mRNA in simple retroviruses^{8,9}. An excess of the CTE inhibits cellular mRNA export^{10,11}, and Tap is thought to be the limiting factor titrated by the CTE⁸.

To determine whether Aly is present in the purified spliced mRNP or Δ i-mRNP, Western analysis was carried out. The Aly antibody recognizes a single band of the correct size (30 kilodaltons)^{12,13} in HeLa nuclear extract (Fig. 2, lanes 1). Strikingly, Aly is highly abundant in the mRNP generated by splicing of adenovirus major late (AdML) pre-mRNA but not in the AdML Δ i-mRNP (Fig. 2a). The same results were obtained with Fushi tarazu (Ftz)-spliced mRNP and Δ i-mRNP (Fig. 2b). In contrast to Aly, heterogeneous nuclear RNP (hnRNP) A1 was detected at similar levels in both the spliced and Δ i-mRNPs (Fig. 2, lower panels). We conclude that Aly is recruited to the spliced mRNP in a splicing-dependent manner and that this recruitment is likely to be general because it occurs with different spliced mRNAs.

It has been reported that Aly is an hnRNP-like protein because it contains an RNA-binding motif, binds to RNA, and is located in the nucleus^{6,14}. However, many proteins, such as the SR (serine-arginine) family of essential splicing factors, have these properties but are not hnRNP proteins¹⁵. To determine whether Aly associates with hnRNP complexes, we assembled the H complex on AdML or Ftz pre-mRNA (Fig. 2, lanes 2). The H complex assembles non-specifically on all RNAs incubated in nuclear extracts and contains the approximately

20 hnRNP proteins that package nascent transcripts *in vivo*^{16,17}. In contrast to hnRNP A1, Aly is barely detectable in either of the H complexes. We conclude that Aly is not an H-complex component and is therefore not likely to be an hnRNP protein.

We next asked when Aly is recruited during spliceosome assembly. Pre-mRNA was incubated in splicing extracts for the indicated times, and the resulting complexes were MBP-affinity purified (Fig. 3). To investigate the strength of Aly association, the complexes were isolated in either low-salt (i.e. functional conditions, 60 mM) or high-salt (250 mM) buffer. Only unspliced pre-mRNA was detected in the spliceosomes after 7 and 14 minutes. After 21 minutes, splicing intermediates and low levels of spliced mRNA were detected. Spliced mRNA accumulates after 40 minutes and is the main RNA species after 90 minutes.

We first carried out Western analysis of the purified complexes using antibodies against the U2 snRNP-specific B'' protein and U5 snRNP-specific 40 K and 116 K proteins^{18–20}. As expected²¹, U2 snRNP is tightly associated with the spliceosome after 7 minutes, but a lower level of U5 snRNP is present (Fig. 3b, c and d). Later, both snRNPs are present. In contrast to the snRNP proteins, the levels of Aly are significantly different throughout the time course and in low- and high-salt conditions (Fig. 3e). In low salt, Aly is first detected in the spliceosome after 7 minutes and is more abundant later. In high salt, Aly is tightly associated only with the complexes containing high levels of spliced mRNA (40 and 90 minutes). We conclude that Aly is recruited during spliceosome assembly and then becomes more tightly associated with the spliced mRNP.

As Tap has been implicated in mRNA export and interacts directly with Aly^{2,6,8,22,23}, we next examined whether Tap was present in the purified complexes. In contrast to Aly, Tap is present in much lower levels in the spliced mRNP than in the nuclear extract, and even less Tap is detected in the high-salt-spliced mRNP (Fig. 3f). Thus, it is possible that Tap is normally recruited efficiently at a later stage in the export pathway. Alternatively, Tap may be bound more weakly to the spliced mRNP than is Aly.

We next investigated whether Aly is an mRNA export factor in metazoans by testing the effects of excess recombinant Aly on export. Short times (45, 90 and 135 minutes) were used for the export assay because the pre-mRNA is spliced at that stage, but little export has occurred (Fig. 4b, + buffer, upper panel)¹. Significantly, glutathione S-transferase (GST)-Aly increases both the rate and

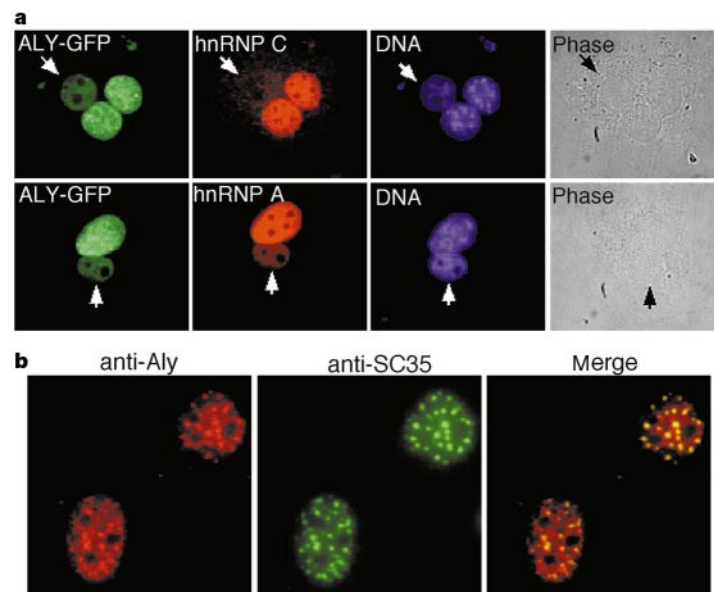


Figure 5 Aly co-localizes with splicing factors in nuclear speckles and shuttles between the nucleus and cytoplasm. **a**, Heterokaryon assays were performed by co-culturing *Xenopus* and mouse cell lines, and hnRNP A1, hnRNP C and Aly were detected. The

arrows indicate *Xenopus* nuclei. **b**, Affinity-purified Aly or SC35 antibodies were used for double indirect immunofluorescence of HeLa cells. GFP, green fluorescent protein.

efficiency of spliced mRNA export (Fig. 4b and c). This stimulation was not observed with any other proteins tested, including hnRNP A1 or GST-Dbp5 (data not shown). The export of tRNA is not affected by excess GST-Aly, indicating that Aly is specific for mRNA export (see Supplementary Information).

In contrast to the spliced mRNA, Δ i-mRNA is not exported efficiently (Fig. 4b and c). Slightly more Δ i-mRNA export is observed in the presence of excess Aly (Fig. 4b and c), and Aly also promotes some export of unspliced pre-mRNA (Fig. 4b). This stimulation may occur because GST-Aly is present in excess. Thus, it is possible that Aly is normally tightly regulated to ensure that only spliced mRNAs are efficiently exported.

To determine whether the stimulation of mRNA export is a direct effect of GST-Aly, we incubated the 90-minute nuclei with glutathione beads, which specifically bind to GST-Aly. Only the spliced mRNA and a lower level of the Δ i-mRNA bind to the beads, with no detectable binding of the intron, U6 snRNA or tRNA (Fig. 4c, lanes 3, 4, 7 and 8, and see Supplementary Information). These data indicate that the stimulation of export is a direct effect of GST-Aly (Fig. 4b). Although the spliced mRNA binds to the GST beads only twofold more efficiently than the Δ i-mRNA, this may be due in part to a low level of indiscriminant binding of excess GST-Aly to the Δ i-mRNA. In addition, the spliced mRNP may not bind as efficiently to the beads because of the large size and/or conformation of this complex. Consistent with this explanation, the spliced mRNP binds to the MBP-affinity resin about 3 times less efficiently than does the Δ i-mRNP.

The data presented above indicate that Aly is an mRNA export factor in metazoans. To determine whether Aly shuttles between the nucleus and cytoplasm, we carried out a heterokaryon shuttling assay. Both Aly and hnRNP A1 are detected in the *Xenopus* and mouse nuclei in the heterokaryon, whereas hnRNP C remains in the mouse nucleus (Fig. 5a). We conclude that Aly shuttles, and thus may accompany the mRNA during export to the cytoplasm.

Essential splicing factors, including snRNP and non-snRNP proteins, are concentrated within about 50 nuclear speckles in mammalian cells²⁴, whereas the approximately 20 hnRNP proteins are uniformly distributed throughout the nucleus¹⁷. As shown by immunofluorescence, Aly is present in nuclear speckles and co-localizes with the essential splicing factor SC35 (Fig. 5b)²⁵. The observation that Aly is present in speckles together with splicing factors is consistent with our data showing that Aly is recruited to mRNA in a splicing-dependent manner. Splicing factors are thought to move from the speckles to sites of pre-mRNA synthesis²⁶. Thus, Aly is compartmentalized with factors destined to associate with spliceosomes. The presence of Aly in speckles, the splicing-dependent recruitment of Aly to mRNA, and the absence of Aly from hnRNP complexes indicate that Aly is not an hnRNP protein.

Here we provide direct evidence that two major nuclear machineries, those for pre-mRNA splicing and mRNA export, are directly coupled. Specifically, Aly is recruited to the pre-mRNA during spliceosome assembly; it becomes tightly bound in the spliced mRNP, and promotes export of mRNA generated by the spliceosome. Aly also shuttles between the nucleus and cytoplasm *in vivo* and co-localizes with splicing factors in nuclear speckles. Thus, Aly not only has the properties expected of an export factor but is also compartmentalized with spliceosome components. It has been proposed that mRNAs may be exported by multiple mechanisms. In particular, different shuttling hnRNP proteins, such as hnRNP A1, have been proposed to bind to different mRNAs in a sequence-dependent manner and mediate mRNA export²⁷. Although there are clearly RNA-binding proteins that are unique to different mRNAs, our data indicate that Aly does not fall into this category. Instead, Aly appears to be a general mRNA export factor and a component of a general mRNA export machinery. In support of this view, splicing promotes export of all mRNAs tested¹, and we have shown here that Aly is recruited to different mRNAs during splicing. Thus,

splicing-dependent recruitment of Aly is likely to be a general phenomenon. Additionally, Aly is a mammalian homologue of Yra1p (refs 2, 6), an essential mRNA export factor in yeast, indicating that Aly is a component of a conserved mRNA export machinery. □

Methods

Plasmids encoding AdML and Ftz derivatives and GST-Aly were described in refs 1 and 4, and in ref. 12, respectively. The spliced mRNP and Δ i-mRNP were isolated by gel filtration followed by MBP-affinity purification as described in the Fig. 1 legend and ref. 4. ³²P-labelled RNA was fractionated on 8% (Fig. 1b and Fig. 4) or 15% (Fig. 1a and Fig. 3a) denaturing polyacrylamide gels. Proteins in MBP-affinity-purified complexes were separated on NuPAGE 4–12% Bis-Tris gels (Invitrogen). Western blots were probed with antibodies against U2 B', U5 116 K, U5 116 K, Tap, and hnRNP A1^{18–20,23,28}. Antibodies against Aly¹³ and the C terminus of Tap²³ were affinity purified against recombinant proteins. Immunofluorescence was carried out as described²³. For the shuttling assay, a mouse fibrosarcoma cell line L929 stably expressing Aly conjugated with green fluorescent protein was co-cultured with *Xenopus* A6 cells in the presence of 100 μ g ml⁻¹ cycloheximide. Cells were fused by PEG1500 and incubated in the presence of cycloheximide for 4 hours^{23,29}. The cells were fixed, permeabilized, and incubated with hnRNP C or hnRNP A1 monoclonal antibodies followed by incubation with cy3-labelled anti-mouse IgG.

Received 15 June; accepted 3 August 2000.

- Luo, M. & Reed, R. Splicing is required for rapid and efficient mRNA export in metazoans. *Proc. Natl. Acad. Sci. USA* **96**, 14937–14942 (1999).
- Strasser, K. & Hurt, E. Yra1p, a conserved nuclear RNA-binding protein, interacts directly with Mex67p and is required for mRNA export. *EMBO J.* **19**, 410–420 (2000).
- Le Hir, H., Moore, M. J. & Maquat, L. E. Pre-mRNA splicing alters mRNP composition: evidence for stable association of proteins at exon-exon junctions. *Genes Dev.* **14**, 1098–1108 (2000).
- Das, R., Zhou, Z. & Reed, R. Functional association of U2 snRNP with the ATP-independent spliceosomal complex E. *Mol. Cell* **5**, 779–787 (2000).
- Hamm, J. & Mattaj, J. W. Monomethylated cap structures facilitate RNA export from the nucleus. *Cell* **63**, 109–118 (1990).
- Stutz, F. *et al.* REF, an evolutionary conserved family of hnRNP-like proteins, interacts with TAP/Mex67p and participates in mRNA nuclear export. *RNA* **6**, 638–650 (2000).
- Segref, A. *et al.* Mex67p, a novel factor for nuclear mRNA export, binds to both poly(A)⁺ RNA and nuclear pores. *EMBO J.* **16**, 3256–3271 (1997).
- Gruter, P. *et al.* TAP, the human homolog of Mex67p, mediates CTE-dependent RNA export from the nucleus. *Mol. Cell Biol.* **1**, 649–659 (1998).
- Ernst, R. K., Bray, M., Rekosh, D. & Hammarström, M. L. A structured retroviral RNA element that mediates nucleocytoplasmic export of intron-containing RNA. *Mol. Cell Biol.* **17**, 135–144 (1997).
- Pasquinelli, A. E. *et al.* The constitutive transport element (CTE) of Mason-Pfizer monkey virus (MPMV) accesses a cellular mRNA export pathway. *EMBO J.* **16**, 7500–7510 (1997).
- Saavedra, C., Felber, B. & Izaurralde, E. The simian retrovirus-1 constitutive transport element, unlike the HIV-1 RRE, uses factors required for cellular mRNA export. *Curr. Biol.* **7**, 619–628 (1997).
- Bruhn, L., Munnerlyn, A. & Grosschedl, R. ALY, a context-dependent coactivator of LEF-1 and AML-1, is required for TCR α enhancer function. *Genes Dev.* **11**, 640–653 (1997).
- Wichmann, I., Garcia-Lozano, J. R., Respaldiz, N., Gonzalez-Escribano, M. F. & Nunez-Roldan, A. Autoantibodies to transcriptional regulation proteins DEK and ALY in a patient with systemic lupus erythematosus. *Hum. Immunol.* **60**, 57–62 (1999).
- Portman, D. S., O'Connor, J. P. & Dreyfuss, G. YRA1, an essential *Saccharomyces cerevisiae* gene, encodes a novel nuclear protein with RNA annealing activity. *RNA* **3**, 527–537 (1997).
- Fu, X. D. The superfamily of arginine/serine-rich splicing factors. *RNA* **1**, 663–680 (1995).
- Bennett, M., Pinol-Roma, S., Staknis, D., Dreyfuss, G. & Reed, R. Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly *in vitro*. *Mol. Cell Biol.* **12**, 3165–3175 (1992).
- Dreyfuss, G., Matunis, M. J., Pinol-Roma, S. & Burd, C. G. hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.* **62**, 289–321 (1993).
- Habets, W. J., Hoet, M. H., De Jong, B. A., Van der Kemp, A. & Van Venrooij, W. J. Mapping of B cell epitopes on small nuclear ribonucleoproteins that react with human autoantibodies as well as with experimentally-induced mouse monoclonal antibodies. *J. Immunol.* **143**, 2560–2566 (1989).
- Achsel, T., Ahrens, K., Brahm, H., Teigelkamp, S. & Luhrmann, R. The human U5-220kD protein (hPrp8) forms a stable RNA-free complex with several U5-specific proteins, including an RNA unwindase, a homologue of ribosomal elongation factor EF-2, and a novel WD-40 protein. *Mol. Cell Biol.* **18**, 6756–6766 (1998).
- Fabrizio, P., Lagerbauer, B., Lauber, J., Lane, W. S. & Luhrmann, R. An evolutionarily conserved U5 snRNP-specific protein is a GTP-binding factor closely related to the ribosomal translocase EF-2. *EMBO J.* **16**, 4092–4106 (1997).
- Staley, J. P. & Guthrie, C. Mechanical devices of the spliceosome: motors, clocks, springs, and things. *Cell* **92**, 315–326 (1998).
- Kang, Y. & Cullen, B. R. The human Tap protein is a nuclear mRNA export factor that contains novel RNA-binding and nucleocytoplasmic transport sequences. *Genes Dev.* **13**, 1126–1139 (1999).
- Katohira, J. *et al.* The Mex67p-mediated nuclear mRNA export pathway is conserved from yeast to human. *EMBO J.* **18**, 2593–2609 (1999).
- Spector, D. L. Macromolecular domains within the cell nucleus. *Annu. Rev. Cell Biol.* **9**, 265–315 (1993).
- Fu, X. D. & Maniatis, T. Factor required for mammalian spliceosome assembly is localized to discrete regions in the nucleus. *Nature* **343**, 437–441 (1990).
- Lewis, J. D. & Tollervey, D. Like attracts like: getting RNA processing together in the nucleus. *Science* **288**, 1385–1389 (2000).
- Nakiely, S. & Dreyfuss, G. Transport of proteins and RNAs in and out of the nucleus. *Cell* **99**, 677–690 (1999).
- Pinol-Roma, S., Choi, Y. D., Matunis, M. J. & Dreyfuss, G. Immunoprecipitation of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes Dev.* **2**, 215–227 (1988); erratum *Genes Dev.* **2**, 490 (1988).

29. Pinol-Roma, S. & Dreyfuss, G. Shuttling of pre-mRNA binding proteins between nucleus and cytoplasm. *Nature* **355**, 730–732 (1992).
 30. Reed, R. Protein composition of mammalian spliceosomes assembled *in vitro*. *Proc. Natl Acad. Sci. USA* **87**, 8031–8035 (1990).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We are indebted to K. N. Clouse for discussions and comments on the manuscript. We thank R. Luhrmann for U5 antibodies, G. Dreyfuss for the hnRNP A1 and C antibodies, and A. Nunez-Roldan for the Aly antibody.

Correspondence and requests for materials should be addressed to R.R. (e-mail: reed@hms.harvard.edu).

Chromodomains are protein–RNA interaction modules

Asifa Akhtar*, Daniele Zink† & Peter B. Becker*

* Adolf Butenandt-Institut, Molekularbiologie, and † Institut für Anthropologie und Humangenetik, Ludwig-Maximilians-Universität, 80336 München, Germany

In *Drosophila*, compensation for the reduced dosage of genes located on the single male X chromosome involves doubling their expression in relation to their counterparts on female X chromosomes¹. Dosage compensation is an epigenetic process involving the specific acetylation of histone H4 at lysine 16 by the histone acetyltransferase MOF^{2–5}. Although MOF is expressed in both sexes, it only associates with the X chromosome in males. Its absence causes male-specific lethality⁶. MOF is part of a chromosome-associated complex comprising male-specific lethal (MSL) proteins and at least one non-coding roX RNA⁷. How MOF is integrated into the dosage compensation complex is unknown. Here we show that association of MOF with the male X chromosome depends on its interaction with RNA. MOF specifically binds through its chromodomain to roX2 RNA *in vivo*. *In vitro* analyses of the MOF and MSL-3 chromodomains indicate that these chromodomains may function as RNA interaction modules. Their interaction with non-coding RNA may target regulators to specific chromosomal sites.

The association of (MSL) proteins (MSL-1, MSL-2, MSL-3, MOF, MLE) and roX RNA with the male X chromosome has been visualized by immunofluorescence analysis of larval polytene chromosomes^{3,8–10}. *Drosophila* SL-2 cells can act as a model system for male features, as functional dosage compensation complex (DCC) has been purified from them^{5,11}, and they have also been used to study the post-transcriptional regulation of dosage compensation¹². The nuclear territory of the X chromosome in SL-2 cells can be visualized by immunostaining with antisera against MSL-1 (Fig. 1a), MOF (Fig. 1c), the histone H4 isoform acetylated at Lys 16 (K16; Fig. 1d) and MLE (see Supplementary Information). In agreement with previous observations¹³, MSL-1 remained localized to the X chromosome after permeabilization of the cells and RNase treatment (Fig. 1, + RNase). In contrast, the bulk of MOF staining disappeared from the X chromosome after RNase treatment (Fig. 1f). MLE, which interacts with larval polytene chromosomes in an RNase-sensitive manner¹³, had dissociated from the chromosome upon permeabilization of the cells (see Supplementary Information); therefore, we could not confirm the RNase sensitivity of MLE's chromosomal association in this system. Loss of MOF correlated with a reduction of the H4 acetyl-K16 histone isoform at

the X chromosome, suggesting a high turnover of the modification under those conditions (Fig. 1d, g). These results suggest that the stable integration of MOF into chromosome-bound DCC involves an RNase-sensitive structure, and that the continued association of MOF with DCC is independent of MLE.

Candidate RNAs that may contribute to the assembly and chromosomal association of the DCC are the roX1 and roX2 RNAs that are stably expressed in male but not female flies, and that colocalize with MSL proteins on the male X chromosome^{14,15}. The two RNAs have no effect on the integrity of DCC, as deletion of a single roX RNA has no phenotype⁸, but mutation of both genes abolishes the interaction of MSL proteins with the chromosome¹⁶. SL-2 cells express only roX2 RNA, which forms a complex with MSL proteins^{5,15} and is therefore an excellent candidate for an RNA involved on the targeting of MOF. To test whether MOF interacts with roX2 RNA *in vivo*, we prepared nuclear extracts from SL-2 cells and isolated the soluble DCC by immunoprecipitation with antibodies specific for either MOF or MLE. We tested the RNase sensitivity of interactions by RNase treatment of extracts before immunoprecipitation. We characterized the immunoprecipitated complexes by western blot analysis. Under these conditions, MOF, MLE, MSL-3 (Fig. 2a) and MSL-2 (data not shown) co-immunoprecipitated independently of RNase treatment, indicating that the DCC was held together by protein–protein interactions and/or that bridging RNA was protected from the nuclease attack. In the absence of RNase, we readily detected roX2 RNA in the immunopurified complex by reverse transcription of RNA followed by polymerase chain reaction (RT–PCR) with roX2 specific primers, whether the immunoprecipitation had been performed with antibodies specific for MOF or MLE (Fig. 2b, lanes 3, 5).

These experiments confirm earlier reports that roX2 RNA can be part of soluble DCC^{5,15}, but they raise the issues of how soluble DCC may differ from the chromosome-associated complex and whether MOF interacts with roX2 RNA directly or indirectly (for example, through MLE¹³). To clarify the latter issue, we made use of the observation that the association of MLE with DCC is sensitive to

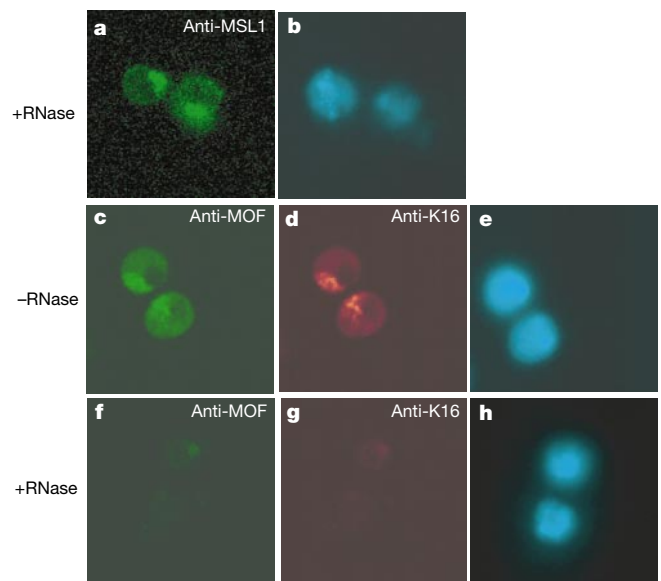


Figure 1 Localization of MOF to the X chromosome is RNase-sensitive. SL-2 cells were detergent-permeabilized, fixed and immunostained with antibodies against MSL-1 (a), MOF (c) and the acetylated Lys 16 of histone H4 (anti-K16) (d). Nuclei in b, e and h were stained with DAPI. Cells in a and f were RNase-treated after permeabilization. The same nuclei were stained with either anti-MOF or anti-K16 in c and d, respectively. The gain (voltage) and the laser power were not changed when untreated and RNase treated cells were scanned for comparison; show the relative intensities of the raw data are shown in c, d, f, g.

elevated ionic strength⁵. Under those stringent immunoprecipitation conditions, MLE was no longer associated with MOF, although MSL-2 and MSL-3 were still detectable (Fig. 2c). In the absence of MLE, a significant amount of roX2 RNA remained associated with MOF (Fig. 2d, lanes 1, 2). The specificity of this interaction was established by showing that an abundant control RNA, encoding heat-shock protein 26 (hsp26), was not pelleted (Fig. 2d).

Although these experiments ruled out that MLE, to date the best candidate for an RNA interacting factor¹³, was bridging between MOF and roX2, the involvement of MSL-3 (see below) or other unknown protein(s) remained possible. To determine whether MOF interacted with RNA directly, we studied the recombinant enzyme *in vitro*⁴. Electrophoretic mobility shift assay (EMSA) with fragments derived from roX1 (data not shown) or hsp26 RNA (Fig. 3a) revealed a nonspecific interaction of MOF with RNA. Competition experiments showed that MOF interacts with RNA with high preference over DNA (Fig. 3a). We previously showed that MOF is unable to interact with DNA efficiently⁴. A lack of specificity of the RNA interaction may be due to misfolding of the *in vitro* transcribed RNAs in the absence of chaperones¹⁷ or, most likely, to the use of arbitrary fragments of roX RNA which lack optimum binding sites.

The specific interaction of MOF with RNA should be distinguished from insignificant 'sticking', by determining whether MOF has a specific domain for RNA interaction. We therefore mapped the RNA interaction domain, creating a series of recombinant MOF derivatives (Fig. 3b) that were all active in chromatin binding and histone acetylation (ref. 4; and data not shown), and analysing their potential to interact with RNA (Fig. 3c). A MOF protein truncated at its amino terminus (Δ N352) still interacted with RNA, and further deletion of the chromodomain (Δ N518) abolished this interaction (Fig. 3c). As this result suggested that the chromodomain was involved in RNA binding, we mutated hydrophobic residues that are conserved in chromodomains from various origins^{18,19}. MOF derivatives with single amino-acid exchanges (W426G and Y416D) were unable to interact with RNA, whereas a point mutation in the acetyl CoA-binding site, G691E⁴, did not affect RNA binding (Fig. 3c). The W426G and Y416D mutations affected the structure of the chromodomain only locally as the mutant enzymes were still active histone acetyltransferases (HATs). We obtained confirmatory results with an alternative binding assay involving roX1 RNA immobilized on streptavidin-sepharose beads (data not shown).

To establish the physiological relevance of the chromodomain for RNA interaction *in vivo* we transiently expressed MOF derivatives

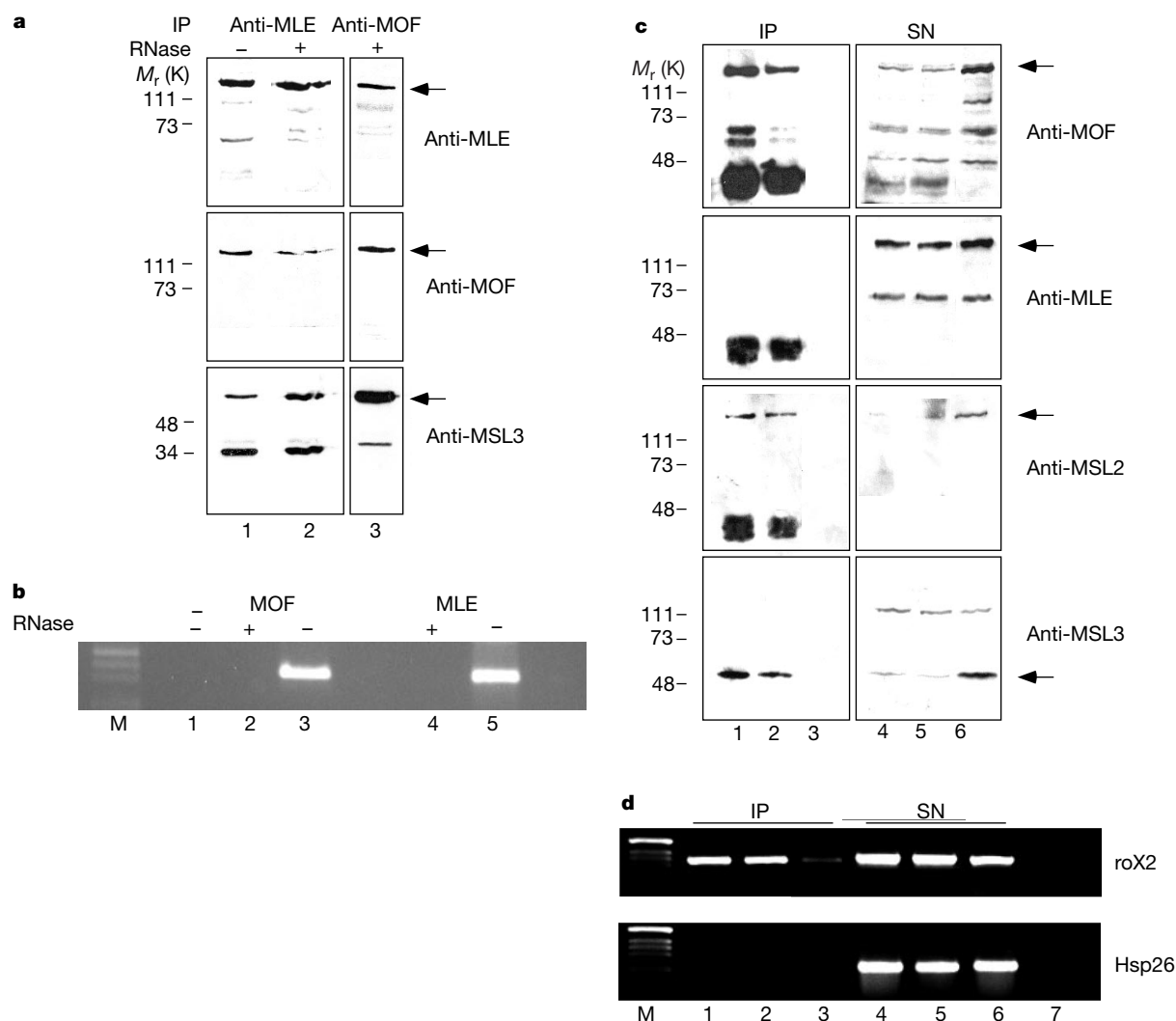


Figure 2 MOF binds roX2 RNA *in vivo*. **a**, Immunoprecipitation (IP) from SL-2 extracts with either anti-MLE or anti-MOF in the presence (+) or absence (-) of RNase. **b**, RT-PCR with roX2 primers. Lane 1, reaction without cDNA; lanes 2–5, IP with anti-MOF or anti-MLE in the presence (+) or absence (-) of RNase, as indicated. **c**, Immunoprecipitation with anti-MOF under stringent conditions. Lanes 1 and 2, washes and IP, respectively, at elevated

salt; lane 3, control without antibody; lanes 4–6, supernatants (SN) of the IP reactions. **d**, RT-PCR analysis of samples described in **c** with primers for either roX2 or hsp26 RNA, as indicated. Size markers are shown to the left (M). Antibodies used are shown; arrows indicate appropriate bands.

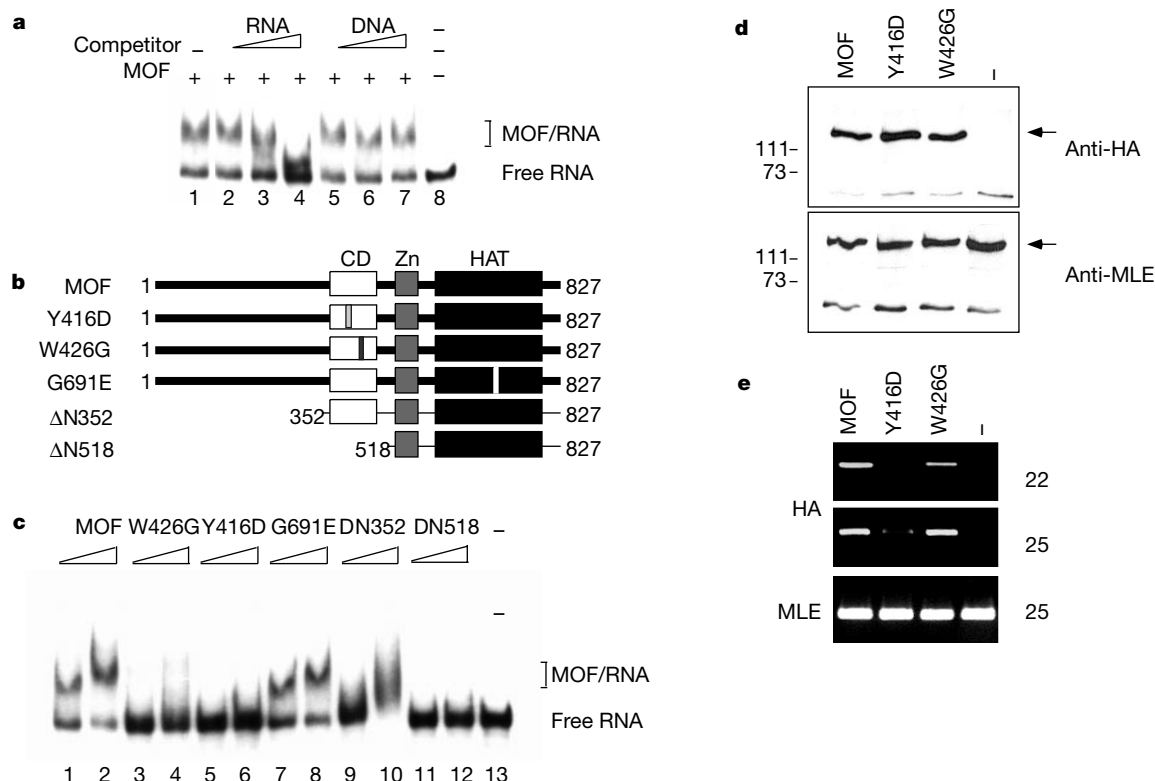


Figure 3 MOF interacts with RNA *in vitro* and *in vivo* through the chromodomain.

a, Electromobility shift assay. Lane 1, interaction of 150 ng MOF with RNA; lanes 2–7, competition with either roX1 RNA (2–4) or DNA (5–7); lane 8, RNA probe.

b, Representation of MOF derivatives. CD, chromodomain; Zn, Zinc finger; HAT, Histone acetyl transferase domain. **c**, Lanes 1–12, EMSA with either 150 ng (odd lanes) or 300 ng

(even lanes) of MOF derivatives as indicated; lane 13, RNA probe. **d**, Extracts from cells transfected with wild-type HA–MOF (MOF), HA–Y416D (Y416D) or HA–W426G (W426G), or untransfected cells (–) were immunoprecipitated with either anti-HA or anti-MLE. Western blots probed with the same antibodies are shown. **e**, RT–PCR of samples in **c** for 22 or 25 cycles (anti-HA), or for 25 cycles (anti-MLE).

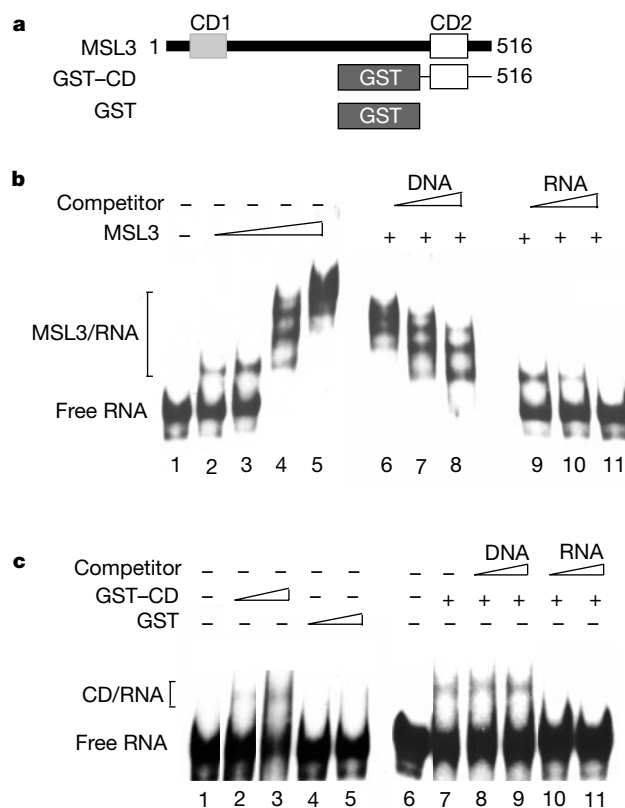


Figure 4 Chromodomain in MSL-3 is also sufficient for interaction with RNA *in vitro*.

a, Representation of MSL-3 derivatives. CD, chromodomain. **b**, Electromobility shift assay. Lane 1, free RNA probe; lanes 2–5, EMSA with 20, 50, 100 or 200 ng of recombinant MSL-3 protein with RNA; lanes 8–11, competition with DNA (6–8) or RNA

(9–11) using 200 ng of MSL-3. **c**, Electromobility shift assay. Lanes 2–5, EMSA with 125 or 250 ng of GST–CD (2, 3) or 250, 500 ng GST (4, 5); lanes 7–11, competition with 0.15 pmol or 0.5 pmol of DNA (8, 9) or RNA (10, 11), or neither (7) using 250 ng of GST–CD; lanes 1, 6, free RNA probe.

from a metallothioneine promoter in SL-2 cells²⁰. The addition of an N-terminal haemagglutinin (HA) epitope allowed us to distinguish ectopic enzymes from endogenous MOF. Whole-cell extracts were prepared before or after induction of transgene expression with copper. Under stringent immunoprecipitation conditions with antibody directed against HA, roX2 RNA, but not hsp26 RNA, associated with HA-MOF (data not shown). When expressed to the same level as intact MOF (Fig. 3d), both chromodomain mutants W426G and Y416D showed an impaired interaction with roX2 RNA. Almost no roX2 RNA could be detected after immunoprecipitation of Y416D, whereas the effect of the W426G mutation was less severe, but still clear (Fig. 3e). We also precipitated MLE, which brought down roX2 RNA, as an internal control for the quality and quantity of the protein extracts (Fig. 3e).

To establish whether RNA interaction is a specific feature of the MOF chromodomain or whether it is a more general property of chromodomains, we tested MSL-3, another dosage compensation protein containing two chromodomains²¹ (Fig. 4a), for interaction with RNA. Recombinant MSL-3 interacted efficiently with RNA, forming several complexes in the EMSA assay at higher concentrations (Fig. 4b, lanes 2–5). Competition experiments established that MSL-3 interacted far better with RNA than with DNA (Fig. 4b, compare lanes 6–8 and 9–11). Even the carboxy-terminal chromodomain of MSL-3 (CD2) when fused to glutathione-S-transferase (GST) was able to interact with RNA but not DNA, whereas the GST moiety alone was inactive (Fig. 4c).

Our results suggest that the MOF chromodomain interacts with roX RNA *in vivo*, which may contribute to the integration of MOF into DCC at the male X chromosome. The association of MOF with DCC is a rather late step in the assembly of the complex and does not occur in the absence of MLE²². The earlier incorporation of roX2 RNA also depends on MLE^{15,16}. The stable association of any one subunit with DCC may rely on multiple interactions with protein and/or RNA subunits, and an additional direct contact between MOF and MLE remains possible. Whether RNA interaction is a general property of chromodomains or restricted to a subfamily of the chromodomain superfamily²¹ remains to be seen. Chromodomains are important for the function of a number of chromatin regulators^{23,24}, but their modes of action have remained enigmatic. Although the related 'chromo shadow domain' of heterochromatin protein 1 (HP1) mediates interactions with several proteins and an interacting peptide has been identified, chromodomains have so far not been shown to contact proteins or peptides²⁵. Mutations of the clr4 protein analogous to the ones that we made in MOF abolish its silencing capacity¹⁸. Small deletions of the polycomb protein containing these residues lead to its delocalization in SL-2 cells²⁶. Non-coding RNAs may be more commonly involved in organizing regulatory complexes than has been appreciated to date^{27,28}. Identification of the RNA structure motif that determines the specific interaction with the chromodomain remains a challenge for the future. Interestingly, dosage compensation in mammals also involves a non-coding RNA, Xist, coating the inactive X²⁹. It is tempting to speculate that roX and Xist RNAs may target regulatory proteins to the X chromosomes of *Drosophila* and humans. □

Methods

Immunofluorescence and confocal microscopy

SL-2 cells were seeded on coverslips at a density of 7×10^6 cells ml⁻¹, fixed with 3.7% formaldehyde for 10 min and blocked with 5% BSA, 0.1% Tween and 0.1% Triton in 4 × SSC (0.66 M NaCl, 0.06 M NaCitrate) at 25 °C for 1 h. They were incubated with primary antibody for 1 h at 25 °C, washed 3 times for 10 min with buffer containing 4 × SSC containing 0.1% Tween and 0.1% Triton, and incubated with fluorescence-conjugated secondary antibody (FITC for experiments except those shown in Fig. 1d and g, where CY3 was used) for 1 h at 25 °C in 4 × SSC containing 5% BSA, 0.1% Tween and 0.1% Triton. We stained the cells with DAPI (0.1 µg ml⁻¹) for 3 min before mounting. For RNase treatment, cells were treated with PBS containing 0.1% Tween/0.1% Triton for 10 min, followed by incubation with either PBS alone or PBS containing 1.0 mg ml⁻¹ RNase A for 10 min, and fixation and staining as above.

Confocal microscopy was performed with a Leica TCS 4D confocal microscope as described³⁰. Images were arranged using Adobe Photoshop 5.0. For double staining, control scans confirmed that no bleed-through was detectable under the conditions used. We obtained antibodies with specificity^{10,13} for MOF, MLE and MSL-1 from M. Kuroda (Baylor College of Medicine, Houston, Texas). No signal was obtained if the first antibody was omitted. The H4Ac16 antiserum was from Upstate Biotechnology.

EMSA and RNA binding assays

For electrophoretic mobility shift assays (EMSA), a 385-base-pair fragment from the hsp26 gene (or a fragment derived from roX1 RNA; data not shown) was transcribed from a linearized plasmid with T7 RNA polymerase, incorporating [³²P]UTP. RNA binding reactions contained 20 mM HEPES pH 7.6, 100 mM KCl, 2 mM EDTA, 0.01% NP40, 1 mM DTT, labelled RNA fragment (10,000 c.p.m.) and 3 µg of transfer RNA. Reactions were incubated on ice for 30 min and electrophoresed on 4% native TBE gels. For competition analysis, 0.05, 0.15 or 0.5 pmol of roX1 RNA or DNA was used unless otherwise indicated.

Extract preparation, immunoprecipitation and RT-PCR

Nuclear extracts were prepared as previously described¹¹. We treated extracts with 250 µg ml⁻¹ RNase A were directly before immunoprecipitation. For immunoprecipitation, 40 µl of extract was incubated with 10 µl polyclonal antibody in buffer A (25 mM HEPES pH 7.6, 150 mM NaCl, 5 mM MgCl₂, 0.1 mM EDTA, 0.1% Tween, 0.2 mg ml⁻¹ BSA) for 2 h at 4 °C, followed by incubation with 50 µl of a 50% (v/v) slurry of protein A agarose beads in buffer A for a further 2 h at 4 °C. After washing the beads in buffer A, the reaction was divided into two. Half of the precipitate was used for protein analysis by western blotting; the other half was used for the amplification of RNA by RT-PCR as described¹⁵. Unless indicated otherwise, 25 cycles of PCR were performed using 15% of the first strand synthesis reaction as template. For stringent immunoprecipitation, the salt in buffer A was increased to 400 mM NaCl. Transfected cells were extracted by lysis in 25 mM HEPES pH 7.6, 100 mM KCl, 12.5 mM MgCl₂, 0.5 mM EDTA, 0.5 mM EGTA, 1 mM DTT, 0.2 mM PMSF, 10% glycerol, 0.5% NP40, by two cycles of freeze/thaw and clearing the lysates by centrifugation.

Transient transfection of SL-2 cells

We carried out transfections with 1 µg of expression clone using the Qiagen Effectene reagent according to the manufacturer's instructions. Forty-eight hours after transfection, cells were induced with 1 mM CuSO₄ for six hours and then collected for extract preparation. N-terminal HA-tagged MOF and mutant derivatives were expressed from the metallothioneine promoter of pRmHa-3 (ref. 20; details are available from the authors on request).

Recombinant proteins

Site-directed mutagenesis and production of the recombinant MSL-3 and MOF derivatives were done as described¹. Residues 437–516 of the MSL-3 protein encompassing chromodomain 2 (CD2, Fig. 3a) were cloned into the pGEX (Pharmacia) vector system and the GST fusion was produced according to manufacturer's instructions. Detailed cloning procedures are available from the authors on request. All clones were verified by sequencing.

Received 21 March; accepted 7 July 2000.

- Lucchesi, J. C. Dosage compensation in flies and worms: the ups and downs of X-chromosome regulation. *Curr. Opin. Genet. Dev.* **8**, 179–184 (1998).
- Turner, B. M. Histone acetylation as an epigenetic determinant of long-term transcriptional competence. *Cell. Mol. Life Sci.* **54**, 21–31 (1998).
- Bone, J. R. *et al.* Acetylated histone H4 on the male X chromosome is associated with dosage compensation in *Drosophila*. *Genes Dev.* **8**, 96–104 (1994).
- Akhtar, A. & Becker, P. B. Activation of transcription through histone H4 acetylation by MOF, an acetyl transferase essential for dosage compensation in *Drosophila*. *Mol. Cell* **5**, 367–375 (2000).
- Smith, E. R. *et al.* The *Drosophila* MSL complex acetylates histone H4 at lysine 16, a chromatin modification linked to dosage. *Mol. Cell. Biol.* **20**, 312–318 (2000).
- Hilfiker, A., Hilfiker-Kleiner, D., Pannuti, A. & Lucchesi, J. C. Mof, a putative acetyl transferase gene related to the Tip60 and MOZ human genes and to the SAS genes of yeast, is required for dosage compensation in *Drosophila*. *EMBO J.* **16**, 2054–2060 (1997).
- Lucchesi, J. C. Dosage compensation: roX marks the spot. *Curr. Biol.* **9**, R807–R808 (1999).
- Meller, V. H., Wu, K. H., Roman, G., Kuroda, M. I. & Davis, R. L. roX1 RNA paints the X chromosome of male *Drosophila* and is regulated by the dosage compensation system. *Cell* **88**, 445–457 (1997).
- Amrein, H. & Axel, R. Genes expressed in neurons of adult male *Drosophila*. *Cell* **88**, 459–469 (1997).
- Kelley, R. L. *et al.* Epigenetic spreading of the *Drosophila* dosage compensation complex from roX RNA genes into flanking chromatin. *Cell* **98**, 513–522 (1999).
- Copp, K. *et al.* Complex formation by the *Drosophila* MSL proteins: role of the MSL-2 RING finger in protein complex assembly. *EMBO J.* **17**, 5409–5417 (1998).
- Merenino, L., Guth, S., Bilbao, D., Martinez, C. & Valcarcel, J. Inhibition of msl-2 splicing by Sex-lethal reveals interaction between U2AF35 and the 3' splice site AG. *Nature* **402**, 838–841 (1999).
- Richter, L., Bone, J. R. & Kuroda, M. I. RNA-dependent association of the *Drosophila* maleless protein with the male X chromosome. *Genes Cells* **1**, 325–336 (1996).
- Stuckenholz, C., Kageyama, Y. & Kuroda, M. I. Guilt by association: non-coding RNAs, chromosome-specific proteins and dosage compensation in *Drosophila*. *Trends Genet.* **15**, 454–458 (1999).
- Meller, V. H. *et al.* Ordered assembly of roX RNAs into MSL complexes on the dosage-compensated X chromosome in *Drosophila*. *Curr. Biol.* **10**, 136–143 (2000).
- Franke, A. & Baker, B. S. The roX1 and roX2 RNAs are essential components of the compensasome, which mediates dosage compensation in *Drosophila*. *Mol. Cell* **4**, 117–122 (1999).

17. Weeks, K. Protein-facilitated RNA folding. *Curr. Opin. Struct. Biol.* **7**, 336–342 (1997).
18. Ivanova, A. V., Bonaduce, M. J., Ivanov, S. V. & Klar, A. J. The chromo and SET domains of the Clr4 protein are essential for silencing in fission yeast. *Nature Genet.* **19**, 192–195 (1998).
19. Ball, L. J. *et al.* Structure of the chromatin binding (chromo) domain from mouse modifier protein. *EMBO J.* **16**, 2473–2481 (1997).
20. Bunch, T. A., Grinblatt, Y. & Goldstein, L. S. Characterization and use of the *Drosophila* metallothionein promoter in cultured *Drosophila melanogaster* cells. *Nucleic Acids Res.* **16**, 1043–1061 (1988).
21. Koonin, E. V., Zhou, S. & Lucchesi, J. C. The chromo superfamily: new members, duplication of the chromo domain and the possible role in delivering transcription regulators to chromatin. *Nucleic Acids Res.* **23**, 4229–4233 (1995).
22. Gu, W., Szaute, P. & Lucchesi, J. C. Targeting of MOF, a putative histone acetyl transferase, to the X chromosome of *Drosophila melanogaster*. *Dev. Genet.* **22**, 56–64 (1998).
23. Pirrotta, V. Polycomb: PcG, TrxG, and chromatin silencing. *Cell* **93**, 333–336 (1998).
24. Jones, D. O., Cowell, I. G. & Singh, P. B. Mammalian chromodomain proteins: their role in genome organisation and expression. *BioEssays* **22**, 124–137 (2000).
25. Smothers, J. F. & Henikoff, S. The HP1 chromo shadow domain binds a consensus sequence pentamer. *Curr. Biol.* **10**, 27–30 (2000).
26. Messmer, S., Franke, A. & Paro, R. Analysis of the functional role of the polycomb chromo domain in *Drosophila melanogaster*. *Genes Dev.* **6**, 1241–1254 (1992).
27. Lanz, R. B. *et al.* A steroid receptor coactivator, SRA, functions as an RNA and is present in an SRC-1 complex. *Cell* **97**, 17–27 (1999).
28. Eddy, S. Noncoding RNA genes. *Curr. Opin. Genet. Dev.* **9**, 695–699 (1999).
29. Panning, B. & Jaenisch, R. RNA and the epigenetic regulation of X chromosome inactivation. *Cell* **93**, 305–308 (1998).
30. Eils, R. *et al.* Three-dimensional reconstruction of painted human interphase chromosomes: active and inactive X chromosome territories have similar volumes but differ in shape and surface structure. *J. Cell. Biol.* **135**, 1427–1440 (1996).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

A.A. acknowledges support from the Human Frontier Science Program. We thank M. Kuroda for providing antibodies; T. Straub for GST control protein; D. Ostareck and G. Mengus for protocols; I. Vetter and C. Schwarzlose for technical assistance; and N. Sadoni for help with arranging images. We are also grateful to W. Hörz, S. Kass and members of the laboratory for critical reading of the manuscript and helpful discussion.

Correspondence and requests for materials should be addressed to P.B.B. (e-mail: pbecker@mol-bio.med.uni-muenchen.de).

Image reconstructions of helical assemblies of the HIV-1 CA protein

Su Li*, Christopher P. Hill†, Wesley I. Sundquist* & John T. Finch†

* Department of Biochemistry, University of Utah, Salt Lake City, Utah 84132, USA

† Division of Structural Studies, MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

The type 1 human immunodeficiency virus (HIV-1) contains a conical capsid comprising ~1,500 CA protein subunits, which organizes the viral RNA genome for uncoating and replication in a new host cell. *In vitro*, CA spontaneously assembles into helical tubes and cones that resemble authentic viral capsids^{1–7}. Here we describe electron cryo-microscopy and image reconstructions of CA tubes from six different helical families. In spite of their polymorphism, all tubes are composed of hexameric rings of CA arranged with approximate local *p6* lattice symmetry. Crystal structures of the two CA domains were 'docked' into the reconstructed density, which showed that the amino-terminal domains form the hexameric rings and the carboxy-terminal dimerization domains connect each ring to six neighbours. We propose a molecular model for the HIV-1 capsid that follows the principles of a fullerene cone⁶, in which the body of the cone is composed of curved hexagonal arrays of CA rings and the ends are closed by inclusion of 12 pentagonal 'defects'.

Although the HIV-1 capsid appears organized, its intrinsic conical asymmetry has prevented molecular structure determinations. Recombinant HIV-1 CA and related proteins encompassing the downstream nucleocapsid (NC) RNA binding domain can spontaneously assemble into both cones⁶ and long helical tubes^{1–7} (Fig. 1a). The cones and tubes seem to be closely related, as they form simultaneously and individual particles can be identified in

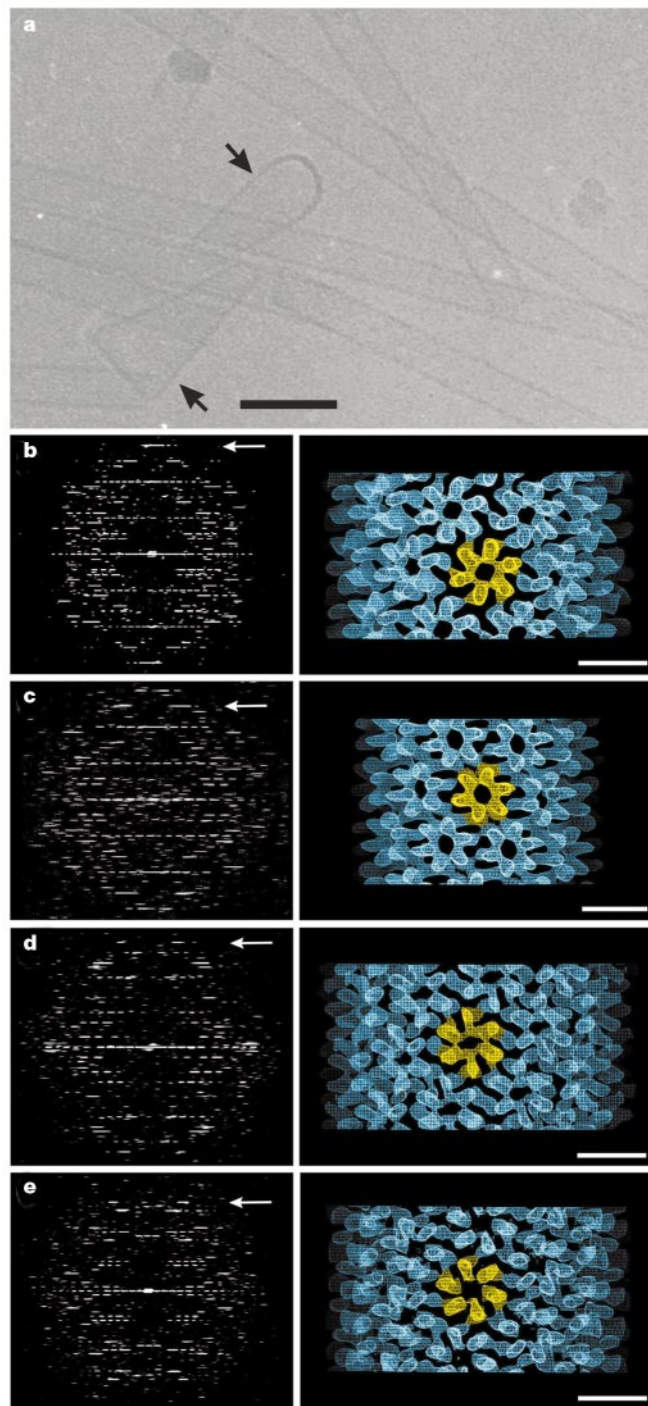


Figure 1 Electron cryo-micrographs and image reconstructions of HIV-1 CA assemblies. **a**, Mixture of tubes and cones (arrows) formed by the purified recombinant HIV-1 CA protein. **b–e**, Computed Fourier transforms (left), and reconstructions (right) of representative CA tubes from the (–12,11) **b**), (–11,10) **c**) (–10,13) **d**) and (–8,5) **e**) helical families. Arrows in the transforms denote layer lines at 0.033 Å^{–1}. The reconstructed densities are contoured at 1.3σ, and individual hexamers in the reconstructions are highlighted for clarity. Scale bars, 100 nm (**a**) and 100 Å (**b–e**).

which a tube turns into a cone at a point of disclination (data not shown). These *in vitro* assemblies appear to be analogues of authentic HIV-1 capsids, because both conical and tubular capsids are also observed in virions (at a ratio of $\sim 20:1$)^{8–10}. Structural analyses of the CA assemblies formed *in vitro* are therefore likely to be relevant for understanding the organization of the viral capsid.

We carried out electron cryo-microscopy (cryo-EM) and image reconstructions to determine the three-dimensional structures of the HIV-1 CA tubes. As the tubes were from a number of distinct helical families, we initially surveyed the range of structures formed by reconstructing nine tubes (from six different families) exhibiting data to at least 30 Å resolution (Table 1). Computed Fourier transforms and three-dimensional reconstructed images of tubes from

four of the different helical families are shown in Fig. 1b–e. All of the helical projections appeared centrosymmetric, and the resulting phase residuals from twofold averaging were low in every case, implying that the surface lattice was based on at least *p2* symmetry, and indicating that the data are of high quality to the reported resolution cut-offs.

Although the computed Fourier transforms differed significantly between the different families, the reconstructed images were highly related. In addition to the local *p2* symmetry, each surface lattice appeared as a network of hexameric rings arrayed with approximate local *p6* symmetry (also reflected in the *p2* unit-cell parameters; Table 1). All subunits comprising the hexameric rings were radially elongated, bi-lobed structures, with their external lobes associating

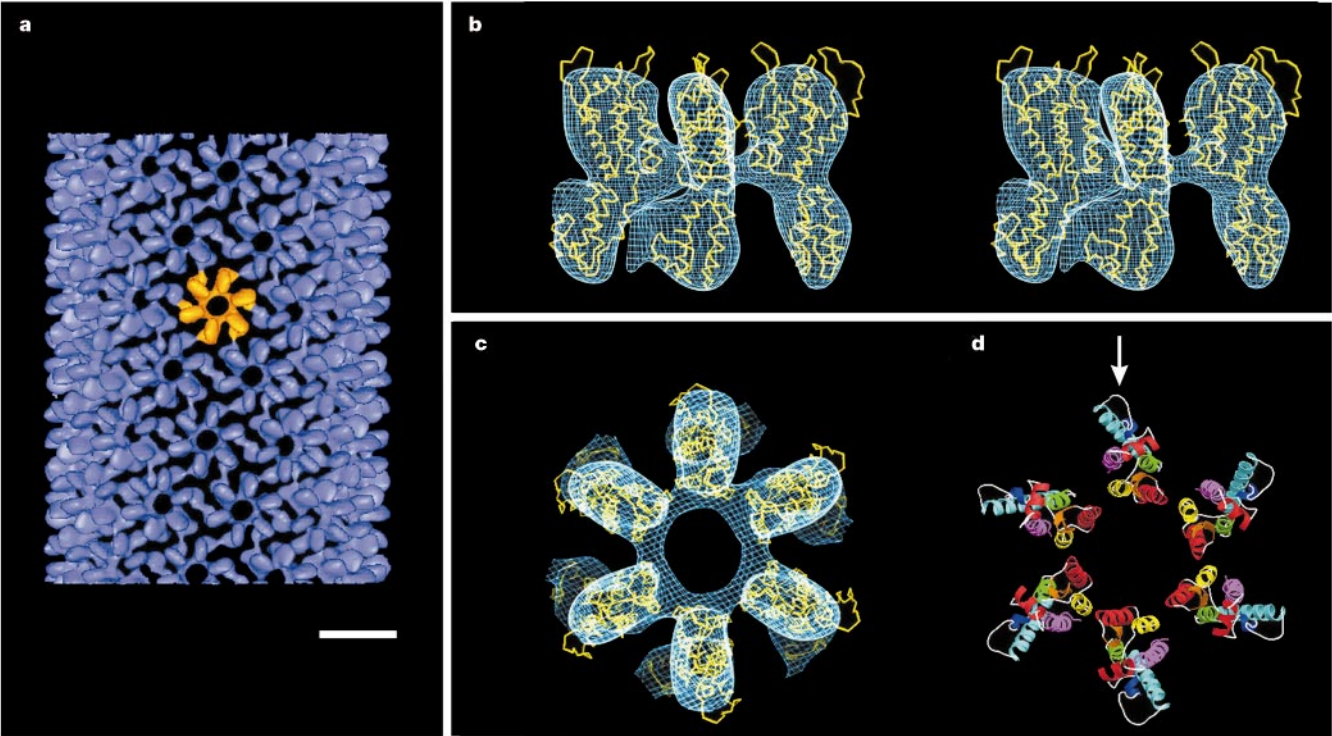


Figure 2 Reconstructed density and molecular modelling of a CA tube. Reconstructions were created by averaging the two highest quality tube images from the (–12,11) family. **a**, View from the tube exterior, with the density contoured at 1.3 σ . A single hexamer is highlighted for clarity. Scale bar, 100 Å. **b**, Stereo image of the CA hexamer viewed with the sixfold axis vertical. Crystal structures of the N- and C-terminal domains of HIV-1 CA were manually ‘docked’ into reconstructed density (1.0 σ) of the three independent CA molecules in the hexamer. The N-terminal domains project towards the top. Although the

N-terminal β -hairpin and cyclophilin-A-binding loops project out of the density here, both loops are highly mobile in solution¹². The optimal docking of the monomeric C-terminal domain is shown. **c**, View of the CA hexamer from the exterior of the tube. The N-terminal CA domains project towards the viewer. **d**, Molecular model for the hexameric ring formed by the N-terminal domain of the HIV-1 CA protein. Structures are coloured as follows: β -hairpin, orange; helix 1, red; helix 2, yellow; helix 3, green; helix 4, cyan; helix 5, dark blue; helix 6, red; helix 7, pink. Arrow denotes a cyclophilin-A-binding loop.

Table 1 Reconstruction data for the HIV-1 CA tubes*

Image	Underfocus† (μm)	$r_{\text{max}}^\ddagger$ (Å)	Helical family§	Resolution cut-off¶ (Å)	Number of layer lines	Phase residual¶¶	<i>p2</i> unit-cell parameters (<i>a</i> , <i>b</i> , γ)#
c421a	1.20	232	(–12,11)	20	46	25.9	106,108,114
c656b	0.55	240	(–12,11)	20	54	28.1	108,111,113
c661a2	2.22	227	(–12,11)	28	24	20.1	106,108,116
9060	1.47	206	(–11,10)	26	22	21.2	100,103,109
c140b	1.46	228	(–11,10)	30	25	17.1	108,110,111
c406b	2.37	263	(–7,9)	30	21	23.2	106,104,121
c419b	1.34	250	(–10,13)	22	37	26.2	115,105,113
c642b	1.17	228	(–8,5)	20	54	28.1	107,109,123
c661b	2.39	247	(–6,8)	30	23	19.0	111,109,123

* Additional data are provided in Supplementary Information.
† Underfocus values were determined from the contrast transfer function using the sector-averaging method.
‡ r_{max} , maximum radius of tube.
§ Helical family designations follow the conventions in ref. 27.
¶ Layer line data were extracted within the first zero of the contrast transfer function.
¶¶ Amplitude-weighted phase deviation from 0° or 180° (Φ_R). $\Phi_R = \Sigma F(R, l) \Delta \Phi(R, l) / \Sigma F(R, l)$; $\Delta \Phi(R, l)$, phase deviation from nearest 0° or 180°; $F(R, l)$, amplitude of the Fourier term at radius *R* on layer line *l*.
p2 unit-cell parameters are calculated at r_{max} . *a* and *b* are in angstroms, γ in degrees.

to form hexamers on the tube exterior. At a lower tube radius, the interior lobes projected down and away from the axes of the hexameric rings with a right-handed aspect, as viewed from the tube exterior, and their densities connected each hexameric ring to its six adjacent neighbours. Thus, the helical lattices appeared as networks of interconnected cogwheels (Fig. 2a).

We averaged the two highest quality images from tubes of the $(-12,11)$ helical family to produce the most reliable reconstruction. The averaging procedure lowered the twofold phase residual (19.6°) and improved the map quality significantly, although basic map features were not altered by any of the averaging steps. The averaged map exhibited high signal to noise and very clear local three- and sixfold symmetry axes (unimposed) (Fig. 2a). This map was therefore used in all subsequent modelling studies. The averaged $(-12,11)$ CA tube was 470 Å in diameter, and the protein shell was ~ 72 Å thick (corresponding to the longest axis of the individual CA subunits). The hexameric rings had exterior diameters of ~ 100 Å, central holes of ~ 25 Å, and inter-ring spacings of ~ 107 Å.

The HIV-1 CA protein comprises two domains separated by a flexible linker sequence. The N-terminal domain is essential for capsid formation, whereas the C-terminal dimerization domain is essential for forming both the immature particle and the mature capsid¹¹. High-resolution structures of both domains have been determined^{12–16}, allowing molecular modelling of the reconstructed CA helix (Fig. 2b–d). The N-terminal domain of CA is an asymmetric, arrowhead shape of about $45 \times 35 \times 18$ Å. This structure fits well into the outer lobe of the reconstructed density, with the longest domain axis roughly perpendicular to the tube axis (Fig. 2b, c).

The N-terminal domains of CA form the hexameric rings, with helices 1 and 2 lining the inner holes of each ring. Intermolecular packing interactions between these two helices presumably stabilize the hexamer, and their polypeptide backbones approach within 8 Å at every interface in our model. The density connecting adjacent subunits is strongest towards the interior of the tube (near the C- and N-terminal ends of helices 1 and 2, respectively). The importance of this region for capsid assembly is also suggested by mutational analyses, which showed that residues in helices 1 and 2 are essential for tube formation *in vitro*, and for replication and capsid formation in cultured virus⁴ (U. K. von Schwedler, K. M. Stray and W.I.S., unpublished data; B. K. Ganser *et al.*, personal communication). Hexamers have not been observed in any of the crystal structures of retroviral CA proteins, and the packing interactions (Fig. 2d) are therefore necessarily different from any previously described crystal lattice interactions. The reconstruction is also incompatible with a different hexamer model that was proposed on the basis of a dimeric lattice contact seen in crystals of the equine infectious anaemia viral CA protein¹⁷. The inner lobes of density that link each hexameric ring to its six nearest neighbours were assigned to the C-terminal domains of CA. The overall size and shape of the monomeric C-terminal CA domain ($\sim 28 \times 28 \times 20$ Å) generally matched the experimental density (Fig. 2b), but unambiguous positioning of this domain will require a higher-resolution reconstruction.

Although structural polymorphism is not uncommon in macromolecular helical assemblies, the degree of polymorphism exhibited by the HIV-1 CA tubes is unprecedented. Three unit-cell vectors define both the $p2$ lattice of a tube's radial projection and its helical family (see Supplementary Information). These vectors also describe the directions of the three helical lines formed by contiguous hexagons within the tube. Reconstructions of the six different families of CA tubes showed that the three unit-cell vectors could adopt essentially any orientation with respect to the helix axis (Fig. 3a). The different tubes also exhibited flexibility in their $p2$ unit-cell dimensions (100 – 115 Å), showing that the packing distance between hexagons can vary. These different types of tube polymorphism were all accommodated primarily by interdomain

'hinge' motions that allowed the C-terminal domains of CA to adopt many orientations with respect to the N-terminal hexameric rings. This flexibility is presumably allowed by the unstructured linker sequence that connects the two domains (CA residues 146 – 149)^{12–16}.

Lentiviral capsids may be assembled on the principles of a fullerene cone^{6,17}, a structure with underlying hexagonal lattice symmetry. The geometric constraints of the hexagonal lattice require that the cones adopt one of five allowed cone angles (19.2° , 38.9° , 60° , 83.6° or 112.9°), and this quantization has been verified experimentally for the CA-NC/RNA cones formed *in vitro*⁶. Our image reconstructions show that the CA tubes are also constructed from curved $p6$ lattices of hexameric rings. We therefore propose that CA cones and tubes are constructed on similar hexagonal nets and differ primarily in the orientations of their hexameric rings. In the tubes the hexamer planes lie parallel to the helix axis, whereas in the cones the hexamers are tilted with respect to the (cone) axis.

We constructed a model for the conical HIV-1 capsid as follows:
(1) The body of the cone is a curved $p6$ surface lattice composed of hexameric CA rings derived from the helical image reconstructions.
(2) The cone is selected to be the narrowest of the allowed fullerene

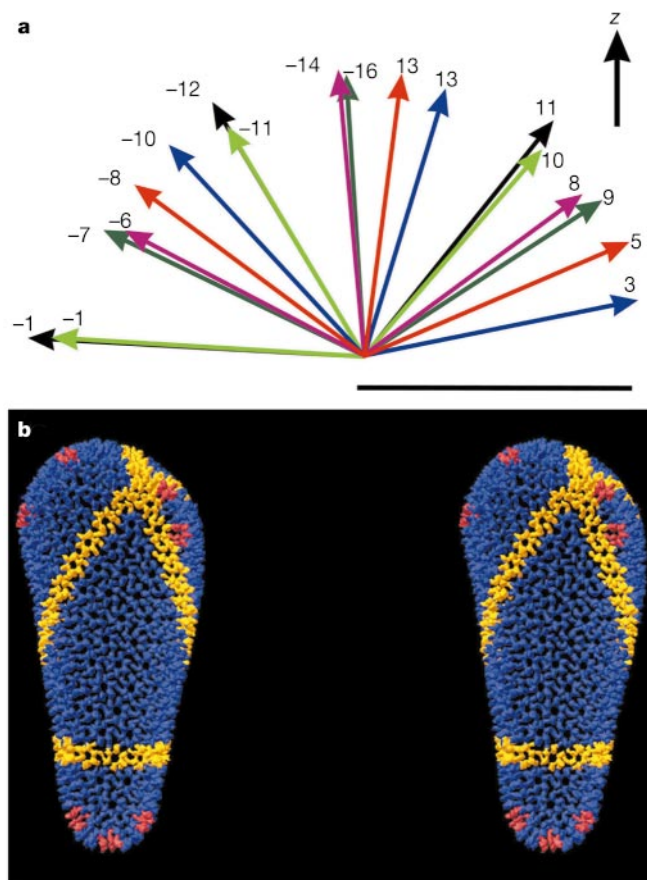


Figure 3 CA helical family polymorphism and cone formation. **a**, Orientations of the unit-cell vectors in the six reconstructed CA helical families drawn to scale and labelled with the corresponding helical start number of contiguous CA hexagons. Tube axes (z) are oriented vertically. Vectors representing the different helical families are coloured as follows: $(-12,11)$, black; $(-11,10)$, light green; $(-7,9)$, dark green; $(-10,13)$, dark blue; $(-8,5)$, red; $(-6,8)$, pink. Scale bar, 100 Å. **b**, Stereo view of a model of the HIV-1 capsid. Pentameric defects are shown in red, and a contiguous line of CA hexamers is highlighted in gold to illustrate how the spiral gradually changes diameter and pitch, and eventually reverses helical hand. Figure created by docking the molecular models shown in Fig. 2 into an idealized fullerene cone of 1,572 CA subunits⁶, and calculating a mask at 8 Å around the C α atoms.

cones (19.2°). Although wider cones are occasionally observed *in vitro*⁶, in SIV⁸ and possibly also in HIV-1 (ref. 10), narrow cones predominate in all cases. (3) The narrow and wide ends of the cone are allowed to close through the introduction of 5 and 7 pentagonal defects, respectively. These 'pentons' are distributed so that the model roughly mimics the shape of an authentic viral capsid, although alternative arrangements are possible at both ends, in analogy to triangulation in icosahedral viruses¹⁸. (4) For simplicity, the pentons are constructed by removing a single subunit from a hexagon and closing the ring (in analogy to strict icosahedral quasi-equivalence), although alternative penton configurations are possible.

As shown in Fig. 3b, contiguous lines of hexagons within the body of the cone trace out spirals that can be described as helices in which the local helical parameters are constantly changing. This model therefore requires that the hexameric CA rings must have the flexibility to adopt many different orientations with respect to the direction and degree of curvature in the cone. This is exactly what is observed in the ensemble of reconstructed CA helices, where the loose packing of adjacent hexameric rings and the hinge between the two CA domains allows great flexibility in the unit-cell orientation and helical family (Fig. 3a). To our knowledge, the long 5,7 fullerene cones formed by the HIV-1 CA protein are unique to the lentiviruses, although aspects of CA assembly are reminiscent of other viral systems that assemble on hexagonal nets (for example, see refs 19–21).

Although many details of the viral capsid structure must await higher resolution studies, some important features are already evident. The CA lattice structure is quite open (Figs 2 and 3), consistent with the idea that the viral capsid primarily organizes, rather than protects, the RNA genome. The open lattice should also allow small macromolecules to diffuse in and out of the viral core, for example allowing nucleotide triphosphates to penetrate and be incorporated by reverse transcriptase *in situ*²².

As described above, the CA subunits are orientated with their N-terminal domains forming the hexameric rings and the C-terminal dimerization domains linking adjacent hexamers. Hexamerization and dimerization might be regulated at different stages of proteolytic maturation and capsid formation^{4,5,23}. In particular, helix 1 shifts significantly upon proteolytic processing at the N-terminal end of CA, and this structural conversion may serve as a morphological switch by promoting hexamerization of the N-terminal domain during viral maturation (ref. 4; and T. L. Stemmler *et al.*, personal communication).

The cellular prolyl isomerase, cyclophilin A (CypA), binds to the HIV-1 CA protein and is packaged within the viral particle at a CypA:CA ratio of about 1:10 (ref. 24). The radial disposition of the N-terminal domains in our model places the CypA-binding loop on the capsid exterior, away from any intermolecular CA interfaces (Fig. 2d). Isolated CypA molecules can be modelled on the capsid surface without steric clash, suggesting that, counter to an earlier proposal¹⁴, low levels of CypA will not promote core disassembly. Saturation of adjacent CypA-binding sites appears sterically disallowed, consistent with the observation that CA tubes can assemble in the presence of low levels of CypA, but not at 1:1 stoichiometries⁷. Although the capsid model does not define the viral function of CypA, our observation that CA mutations that alleviate the viral requirement for CypA also facilitate CA assembly *in vitro* is consistent with a postulated role for the protein as a capsid assembly chaperone⁷.

The interior surface of the capsid model is mainly defined by the highly conserved 20 amino-acid 'major homology region' (MHR) motif that lies at the 'bottom' of the C-terminal domain of CA¹⁶. The precise orientation of this domain is not defined unambiguously, but the MHR motif does not make important intermolecular contacts in any of our models, leaving the polypeptide free to interact with other molecules. This observation is consistent with

possible roles for the MHR in helping to organize genomic and/or reverse transcription complexes for early stages in the retroviral life cycle²⁵.

Although the HIV-1 capsid is highly organized, it is an intrinsically asymmetric object in which every subunit resides in a slightly different environment and the cone length is not uniquely defined. This lack of regularity may pose particular challenges for capsid assembly and may explain why the capsid is assembled through a temporally defined proteolytic pathway. Conversely, the ability to assemble capsids with a range of different volumes may provide the virus with added flexibility in packaging and replicating genomes of different lengths, thereby allowing retroviruses to incorporate cellular genes when advantageous. □

Methods

Sample preparation and data collection

HIV-1 CA was expressed, purified and assembled by incubation of 0.4 mM CA in 1 M NaCl 50 mM Tris-HCl (pH 8.0) for 1 h at 37 °C (ref. 4). Freshly purified CA predominantly formed tubes under all conditions tested. We used wild-type HIV-1_{NL4-3} CA and two proteins with point mutations (A92E, G94D). The two mutations allow viral replication in culture, but eliminate the requirement for the cellular protein cyclophilin A (ref. 26), a putative chaperone for viral capsid assembly⁷. The A92E and G94D proteins formed longer and more regular tubes than wild-type CA, resulting in better quality images. The CA A92E protein was therefore used to generate all of the images described here (but reconstructions of wild-type and G94D CA tubes produced similar density maps; data not shown). Assembly reactions were mixed 1:1 with 0.1 M NaCl on glow-discharged holey carbon grids, blotted and flash-frozen in liquid ethane. EM data were collected on Philips 420 or CM12 electron microscopes at 120 kV under low dose conditions (~10 electrons per Å²) with underfocus values of 1–2 µm. Images were recorded on Kodak SO163 film at nominal magnifications of ×34,000 and ×40,000.

Image analysis and reconstruction

We analysed EM images as described²⁷ using programs developed by N. Unwin (personal communication). Micrographs were digitized using a Zeiss SCAI densitometer (7-µm step size), corresponding to 1.75 Å on the specimen (×40,000). Underfocus and astigmatism values for images were obtained using the sector-averaging method (assuming amplitude contrast of 7%). The repeat distance of each tube was estimated by correlating two separate stretches of tube along its length. Tubes were boxed using this axial repeat length, and the box stretched to 4,096 pixels by interpolating the data in real space. After Fourier transforming (box size 4,096 × 2,048 pixels), this procedure placed all layer lines exactly on the transform grids. Reciprocal lattices were then manually constructed on the computed Fourier transforms. Probable values for the Bessel orders (*n*) for each layer line were estimated. The possible indexing schemes were run in the programs HLXS and SRCH, which gave a clearly lower residual for the correct scheme. Out of plane tilt angles (ω) and x-shift corrections (SRCH output) were input into the program HLXFL, which extracted the layer line data to the resolution cut-off. Phase values were very close to 0° or 180°, indicating that the projected images were centrosymmetric (in good agreement with Fourier filtered images of flattened tubes). We therefore used the program H2FLD for twofold averaging, and the low phase residuals (below 29° in all cases) confirmed the validity of the averaging procedure.

The twofold averaged data were used to calculate final three-dimensional maps using the programs LTLG and HHOR. Intermediate maps all showed similar hexagonal surface lattices, which did not vary significantly with different equator weights (a weight of 0.3 was used). Absolute tilt angles (ω) were determined for the images c656b and c642a by measuring contrast transfer function values at both ends of each tube. The difference between these two values reflected the sense of the out-of-plane tilt angle of the tube (ω), and was used to determine the absolute hand of the reconstructed helix. The local hand of the hexameric rings was the same in both structures, and the hands of the other tubes were adjusted accordingly. As an additional test of the validity of the reconstruction procedure, independent reconstructions of images 9060 and c642b were performed using the MRC helical program package²⁸. The resulting images agreed with those obtained as described above.

We occasionally identified CA (and CA-p2-NC/RNA) tubes that had flattened to produce extended two-dimensional crystals. Fourier filtering showed that both types of projected lattices were composed of six-membered rings of density arrayed with approximate *p6* plane group symmetry, in excellent agreement with the three-dimensional reconstructions. Representative examples of the unit-cell dimensions: CA tube, *a* = 105 Å, *b* = 106 Å, β = 115°; CA-p2-NC/RNA tube, *a* = 90 Å, *b* = 90 Å, β = 120.5°.

Image averaging and model building

Using image c421a as reference, the c656b image was shifted to the same phase origin as c421a, and re-indexed using the program HREINDEX. The two data sets were then merged using HLXADD, and the data from each tube was weighted based upon differences in their repeat lengths. The overall twofold phase residual was 19.6° (34.8° in the 20–25 Å resolution shell). A total of 6,732 CA molecules were averaged to produce the final map. Crystal structures of HIV-1 CA^{14,16} were manually docked into the reconstructed density

using the program O²⁹. We used a single CA molecule as the docking unit; the linker between the N-terminal domain and C-terminal domain (residues 146–149) was allowed full flexibility within steric constraints.

Received 22 February; accepted 5 July 2000.

- Ehrlich, L. S., Agresta, B. E. & Carter, C. A. Assembly of recombinant human immunodeficiency virus type 1 capsid protein *in vitro*. *J. Virol.* **66**, 4874–4883 (1992).
- Campbell, S. & Vogt, V. M. Self-assembly *in vitro* of purified CA-NC proteins from rous sarcoma virus and human immunodeficiency virus type 1. *J. Virol.* **69**, 6487–6497 (1995).
- Groß, I., Hohenberg, H. & Kräusslich, H.-G. *In vitro* assembly properties of purified bacterially expressed capsid proteins of human immunodeficiency virus. *Eur. J. Biochem.* **249**, 592–600 (1997).
- von Schwedler, U. K. *et al.* Proteolytic refolding of the HIV-1 capsid protein amino-terminus facilitates viral core assembly. *EMBO J.* **17**, 1555–1568 (1998).
- Groß, I., Hohenberg, H., Huckhagel, C. & Kräusslich, H.-G. N-terminal extension of human immunodeficiency virus capsid protein converts the *in vitro* assembly phenotype from tubular to spherical particles. *J. Virol.* **72**, 4798–4810 (1998).
- Ganser, B. K., Li, S., Klishko, V. Y., Finch, J. T. & Sundquist, W. I. Assembly and analysis of conical models for the HIV-1 core. *Science* **283**, 80–83 (1999).
- Grattinger, M. *et al.* *In vitro* assembly properties of wild-type and cyclophilin-binding defective human immunodeficiency virus capsid proteins in the presence and absence of cyclophilin A. *Virology* **257**, 247–60 (1999).
- Fukui, T., Imura, S., Goto, T. & Nakai, M. Inner architecture of human and simian immunodeficiency viruses. *Micro. Res. Tech.* **25**, 335–340 (1993).
- Kotov, A., Zhou, J., Flicker, P. & Aiken, C. Association of nef with the human immunodeficiency virus type 1 core. *J. Virol.* **73**, 8824–30 (1999).
- Welker, R., Hohenberg, H., Tessmer, U., Huckhagel, C. & Kräusslich, H. G. Biochemical and structural analysis of isolated mature cores of human immunodeficiency virus type 1. *J. Virol.* **74**, 1168–1177 (2000).
- Dorfman, T., Bukovsky, A., Öhagen, Å., Höglund, S. & Göttinger, H. G. Functional domains of the capsid protein of human immunodeficiency virus type 1. *J. Virol.* **68**, 8180–8187 (1994).
- Gitti, R. K. *et al.* Structure of the amino-terminal core domain of the HIV-1 capsid protein. *Science* **273**, 231–235 (1996).
- Momany, C. *et al.* Crystal structure of dimeric HIV-1 capsid protein. *Nature Struct. Biol.* **3**, 763–770 (1996).
- Gamble, T. R. *et al.* Crystal structure of human cyclophilin A bound to the amino-terminal domain of HIV-1 capsid. *Cell* **87**, 1285–1294 (1996).
- Berthet-Colominas, C. *et al.* Head-to-tail dimers and interdomain flexibility revealed by the crystal structure of HIV-1 capsid protein (p24) complexed with a monoclonal antibody Fab. *EMBO J.* **18**, 1124–1136 (1999).
- Worthylake, D. K., Wang, H., Yoo, S., Sundquist, W. I. & Hill, C. P. Structures of the HIV-1 capsid protein dimerization domain at 2.6 Å resolution. *Acta Crystallogr. D* **55**, 85–92 (1999).

- Jin, Z., Jin, L., Peterson, D. L. & Lawson, C. L. Model for lentivirus capsid core assembly based on crystal dimers of ELAV p26. *J. Mol. Biol.* **286**, 83–93 (1999).
- Caspar, D. & Klug, A. Physical principles in the construction of regular viruses. *Cold Spring Harb. Symp. Quant. Biol.* **27**, 1–24 (1962).
- Katsura, I. Structure and inherent properties of the bacteriophage lambda head shell. I. Polyheads produced by two defective mutants in the major head protein. *J. Mol. Biol.* **121**, 71–93 (1978).
- Steven, A. C. & Trus, B. L. in *Electron Microscopy of Proteins 5*, Viral Structure 1–35 (Academic, London, 1986).
- Cremers, A. M. F., Oostergetel, G. T., Schilstra, M. J. & Mellema, J. E. An electron microscopic investigation of the structure of alfalfa mosaic virus. *J. Mol. Biol.* **145**, 545–561 (1981).
- Zhang, H., Dornadula, G., Alur, P., Laughlin, M. A. & Pomerantz, R. J. Amphipathic domains in the C terminus of the transmembrane protein (gp41) permeabilize HIV-1 virions: a molecular mechanism underlying natural endogenous reverse transcription. *Proc. Natl Acad. USA* **93**, 12519–12524 (1996).
- Groß, I. *et al.* A conformational switch controlling HIV-1 morphogenesis. *EMBO J.* **19**, 103–113 (2000).
- Franke, E. K. & Luban, J. Cyclophilin and Gag in HIV-1 replication and pathogenesis. *Adv. Exp. Med. Biol.* **374**, 217–28 (1995).
- Mammano, F., Öhagen, Å., Höglund, S. & Göttinger, H. G. Role of the major homology region of HIV-1 in virion morphogenesis. *J. Virol.* **68**, 4927–4936 (1994).
- Aberham, C., Weber, S. & Phares, W. Spontaneous mutations in the human immunodeficiency virus type 1 gag gene that affect viral replication in the presence of cyclosporins. *J. Virol.* **70**, 3536–3544 (1996).
- Toyoshima, C. & Unwin, N. Three-dimensional structure of the acetylcholine receptor by cryoelectron microscopy and helical image reconstruction. *J. Cell Biol.* **111**, 2623–2635 (1990).
- Crowther, R. A., Henderson, R. & Smith, J. M. MRC image processing programs. *J. Struct. Res.* **116**, 9–16 (1996).
- Jones, T. A., Zou, J. Y., Cowan, S. W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and location of errors in these models. *Acta Crystallogr. A* **47**, 110–119 (1991).

Supplementary information is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from the London editorial office of Nature.

Acknowledgements

We thank W. Clemons, B. Ganser, J. Smith, C. Villa and D. Worthylake for technical assistance; J. Berriman for help and advice with data collection; and M. Stowell and N. Unwin for guidance in the use of their image reconstruction software. We also thank N. Unwin and R. Henderson for critical reading of our manuscript; and H.-G. Kräusslich for communicating results before publication. This research was supported by the MRC (UK) and by the NIH (US).

Correspondence and requests for materials should be addressed to W.I.S. (e-mail: wes.sundquist@hsc.utah.edu) or J.T.F. (e-mail: jtf1@mrc-lmb.cam.ac.uk).

correction

Signalling through CD30 protects against autoimmune diabetes mediated by CD8 T cells

Christian Kurts, Francis R. Carbone, Matthew F. Krummel, Karl M. Koch, Jacques F. A. P. Miller & William R. Heath

Nature **398**, 341–344 (1999).

It has come to our attention that the conclusions drawn in this paper need to be revised. We have found that the effects attributed to CD30 were the result of background gene(s) of 129 origin. Subsequent backcrossing of the CD30-deficient OT-I line by G. Davey in our laboratory segregated the capacity of low numbers of OT-I cells to cause disease from their capacity to express CD30. Our conclusion that signalling through CD30 protects against autoimmune diabetes must be reconsidered as an effect of a background gene(s) that we are currently mapping. □

erratum

How self-tolerance and the immunosuppressive drug FK506 prevent B-cell mitogenesis

Richard Glynn, Srinivas Akkaraju, James I. Healy, Jane Rayner, Christopher C. Goodnow & David H. Mack

Nature **403**, 672–676 (2000).

The fourth sentence in the Methods section should have read “For stimulation experiments, HEL (Sigma) was used at 500 ng ml⁻¹, goat anti-mu (Jackson Labs) at 10 µg ml⁻¹, FK506 (gift from G. Crabtree laboratory) at 10 ng ml⁻¹, PD98059 (NEB) at 20 µM unless otherwise stated, ionomycin (gift from G. Crabtree laboratory) at 1 µM and EGTA at 3 mM.” □



Cover illustrations courtesy of ITRI, IBM and TSMC. On the right, the first set of four Chinese characters means 'Taiwan science and technology' and the second set of six characters can be translated as 'which way to go?'

Nature London
Porters South, 4 Crinan St,
London N1 9XW, UK
Tel +44 207 833 4000
Fax +44 207 843 4596/7
e-mail: nature@nature.com
http://www.nature.com

Nature Washington
968 National Press Building,
529 14th St NW,
Washington DC 20045, USA
Tel +1 202 737 2355
Fax +1 202 628 1609
e-mail: nature@naturedc.com
http://www.nature.com

Nature Tokyo
Shin-Mitsuke Building (4F),
3-6 Ichigaya Tamachi,
Shinjuku-ku,
Tokyo 162, Japan
Tel +81 3 3267 8751
Fax +81 3 3267 8746
e-mail: nature@naturejpn.com
http://www.naturejpn.com



NATURE VOL 407 21 SEPTEMBER 2000 www.nature.com

regional nature insight

Taiwan

By tapping a reservoir of ethnic Chinese scientific talent in the United States, the little island of Taiwan has achieved a global presence in the electronics industry that far exceeds its geographical size. Chinese scientists returned from the West have also created pockets of excellence in basic research in Taiwan that are beginning to make their mark on the world stage. But Taiwan's very success is placing stress on the island's nascent research infrastructure. The booming electronics industry is sucking some of the best researchers out of the still fragile public sector research system as the flow of returnee scientists from the United States diminishes with the expansion of the US economy. A substantial and growing pool of private and government investment funds is available to support development of new science-based industries in areas such as genomics, biochips and pharmaceuticals. But Taiwan still lacks the fundamental infrastructure to build industries in these fields. So Taiwan is once again turning to its network of Chinese scientists in the United States in search of experienced manpower and promising technologies to bring to the island and co-develop with US partners. Fledgling investment links with the Chinese mainland also hint of a possible dynamic interaction between Chinese scientists in Taiwan, the United States and mainland China that could help bring Taiwan onto the map in biotechnology if the winds of political fortune allow.

David Swinbanks Asia-Pacific Editor
David Cyranoski Asian-Pacific Correspondent

Insight Editor & Publisher: David Swinbanks
Editor, Nature: Philip Campbell
Subeditor: Colin Sullivan
Production Editor: Simon Gribbin
Picture Researcher: Barbara Izdebska
Art Director: Majo Xeridat
Production Manager: Yvonne Strong
Advertisement Manager: Kate Cowan

- 417** Taiwan backs experience in quest for biotech success
- 418** Stepping out from the semiconductor shadow
- 419** Building a progressive sci-tech island
- 420** A science giant rises out of the rice fields
- 421** Nicky Lu—entrepreneur
- 422** Beating the odds
- 424** Homeless, but heading for success
- 425** The rodent route to celebrity
- 426** Bridging the gap

Note from the Editor

Choosing appropriate words when writing about Taiwan is not easy. Is Taiwan a country or part of China? Are the people in Taiwan, Chinese or Taiwanese? Immigration officials stamp 'Republic of China' on the passports of entrants to Taiwan, and officials in Beijing insist unequivocally that Taiwan is part of China. But the recent election of Taiwan's new president Chen Shui-bian, who heads a party containing radicals advocating independence, has raised the political stakes in the highly sensitive issue of Taiwan's status. In this *Regional Insight*, we use the term 'Chinese' to refer to a broad ethnic group of people whether they come from Taiwan, Singapore, Hong Kong, California or mainland China. But where distinction is necessary, we use Taiwanese. The other headache is name order. As all scientists and government officials when dealing with western publications use a western order for their names with surname last, we have respected their wishes. But we have not tried to force a consistent standard by, for example, changing Mao Tse Tung to Tse-Tung Mao, because the traditional Chinese order of his name with surname first is that by which he is well known throughout the world. The same is true of Chen Shui-bian.

Nature® (ISSN 0028-0836) is published weekly on Thursday, except the last week in December, by Nature Publishing Group (Porters South, 4 Crinan Street, London N1 9XW). Registered as a newspaper at the British Post Office. Annual subscription for the Americas US\$595 (institutional/corporate), US\$159 (individual making personal payment), Canada residents please add 7% GST (No. 140911595). North and South American orders to: Nature, Subscription Dept, P.O. Box 5055, Brentwood, TN 37024-5055, USA. Other orders to Nature, Brunel Road, Basingstoke, Hants RG21 2XS, UK. Periodicals postage paid at New York, NY 10010-1707, and additional mailing offices. Authorization to photocopy material for internal or personal use, or internal or personal use of specific clients, is granted by Nature to libraries and others registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided the base fee of \$12.00 an article (or \$2.00 a page) is paid direct to CCC, 222 Rosewood Drive, Danvers, MA 01923, USA. Identification code for Nature: 0028-0836/00 \$12.00+\$2.00. US Postmaster send address changes to: Nature, PO Box 5055, Brentwood, TN 37024-5055. Published in Japan by Nature Japan K.K., Shin-Mitsuke Bldg, 36 Ichigaya Tamachi, Shinjuku-ku, Tokyo 162, Japan. © 2000 Nature Publishing Group.

Taiwan backs experience in quest for biotech success

In under 30 years the tiny island of Taiwan has gone from agricultural backwater to global electronics giant. Now it is turning its attention to biotechnology. But can Taiwan sustain the pace of development?

David Swinbanks and
David Cyranoski

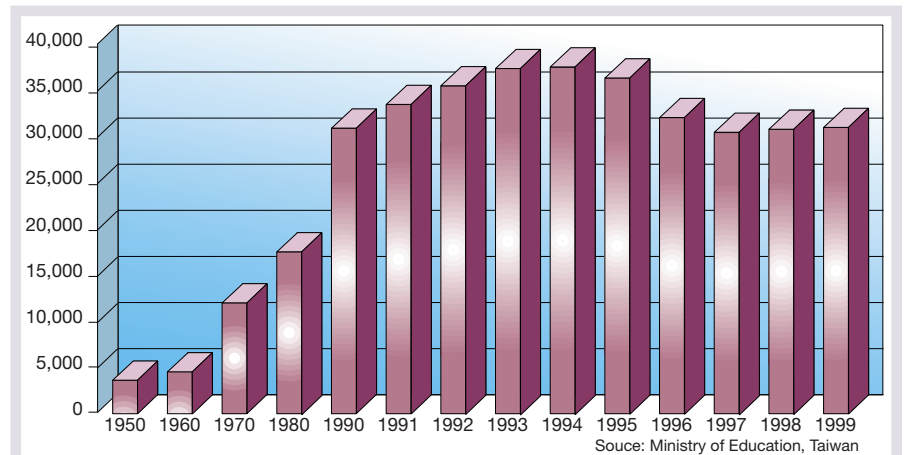
The rise of Taiwan over the past two decades has been meteoric. In the 1960s, it was a poor developing economy with no significant research infrastructure and impoverished universities. Now, the island of 23 million is a global player in the semiconductor industry. It is home to the world's two largest silicon wafer foundries and numerous dynamic start-up companies producing some of the most advanced designs for chips. There are also several star companies in computer manufacture, telecommunications and optoelectronics.

The Hsinchu Science-based Industrial Park (HSIP), home to many of Taiwan's top companies, occupies the space of a small town with just over 80,000 employees. It had annual sales last year of NT\$561 billion (about US\$18 billion) — over 6% of Taiwan's total GDP and accounts for 9% of Taiwan's trade (see Table 1).

The quality of the university system and level of graduate education has improved dramatically, with the National Taiwan University (NTU) in Taipei leading the way. Taiwan now ranks 19th in the world in the Science Citation Index, up from 35th in 1986. And it is ahead of Japan, South Korea and mainland China in terms of the proportion of papers published in six leading scientific journals in each of physics, chemistry and mathematics (see *Nature* 388, 9; 1997).

The Silicon Valley connection

Strong links with Silicon Valley have been vital to Taiwan's success. From the 1950s onwards, Taiwanese students, many of them with degrees in engineering and physics,



Distance learning: the number of Taiwanese students in the United States between 1950 and 1999.

headed for the United States. The brightest of them took research positions in top US universities, in particular Stanford University on the west coast and at the heart of Silicon Valley. The number of Taiwanese students in the United States, along with compatriots from the Chinese mainland, soared in the 1980s peaking in the early 1990s at just under 40,000 (see graph, above).

Astute officials in the Taiwanese government, many of them also educated in the United States, realized the potential of these overseas students. Over the past three decades they have concentrated on creating the right environment to lure back many of these researchers, even those in well established positions at top US companies.

The first of these initiatives was the 1973 merger of several older institutes to form the Hsinchu-based Industrial Technology Research Institute (ITRI). Backed by the



Morris Chang: a semiconductor success story.

powerful Ministry of Economic Affairs, ITRI focuses on semiconductors and computers. This was followed by the creation of the government-backed HSIP in 1980.

In the late 1980s and the first half of the 1990s, thousands of Chinese researchers in Silicon Valley and other parts of the United States chose to return to Taiwan. They came to work at ITRI, which now has over 6,000 personnel, and the HSIP nearby, which has since become the heart of Taiwan's electronics industry.

By the mid-1980s, Taiwan had managed to attract many top Chinese researchers, such as Morris Chang, a Stanford PhD from mainland China. After working his way up to the top ranks of Texas Instruments in Dallas researching germanium and silicon transistors, Chang left his newly acquired position of president and chief executive officer of General Instrument to head ITRI in 1985. Two years later, he founded the Taiwan Semiconductor Manufacturing Company (TSMC) at the HSIP, which is now the world's largest semiconductor foundry.

By the early to mid-1990s, aided by a squeeze on the US economy and restructuring of companies such as IBM and AT&T

Table 1 Status of industry in Hsinchu Science-based Industrial Park (1999)

Industry	No. of companies	No. of employees	Sales (NT\$, bn)	Growth (%)
Integrated circuits	118	48,284	360.8	56
Computers and peripherals	51	16,529	200.9	26
Telecommunications	47	5,299	32.4	22
Optoelectronics	48	11,066	51.4	52
Precision machinery	13	1,165	4.8	43
Biotechnology	15	435	0.7	17
Total	292	82,778	561.0	43

Source: National Science Council, Taiwan; US\$1 = NT\$31

Stepping out from the semiconductor shadow

Success has left Taiwan's Industrial Technology Research Institute (ITRI) with a problem — what to do next. Having helped to revolutionize the country's semiconductor industry, the institute now needs a fresh focus.

Taiwan's premier industrial research institute, ITRI was established in 1973. It is a non-profit organization that helps to promote Taiwanese industry. It has 6,000 researchers, and its 11 laboratories account for more than 500 patents per year.

In the past four years, ITRI has spawned 100 incubators to develop new technology, ten of which have 'graduated' as companies to the Hsinchu Science-based Industrial Park. The institute's Opto-Electronics and Systems laboratories (OES) has helped give Taiwan a 70% market share in the global CD-ROM and DVD-ROM market. ITRI has also established offices overseas in the United States, western Europe, Moscow and Tokyo.

Which way now?

But despite this success, many researchers feel that ITRI needs to redefine its goals. "ITRI was helpful in optoelectronics because industry started late. In other areas — materials, electronics, semiconductors — industry has moved ahead. ITRI needs a structural change," says Yuen-ron Shen at the physics department at the University of California at Berkeley.

Li-chung Lee of ITRI's Silicon Valley office, agrees. "ITRI aims to help small and medium companies that don't have the critical R&D mass to compete with an IBM, but companies are doing more and more R&D themselves. Lately ITRI is doing too much development work," he says.

"ITRI declared victory in [semiconductor manufacture] and then moved on," says Morris Chang, president of ITRI from 1985 until 1987, when he left to form the Taiwan Semiconductor Manufacturing Company. But

where will it go from here?

Geographically, it is stretching out to the new Tainan Science-based Industrial Park in the south. Its overseas offices coordinate joint research efforts with laboratories in their respective regions. "International collaboration and technology transfer will allow ITRI to move away from development to frontier research," says Lee.

ITRI has also been courted by US states that want to develop R&D ties with Taiwan, Lee says, referring specifically to overtures from the office of New York governor George Pataki.

Cashing in on chips

One key organizational change is a concentrated push into biotechnology, with the creation of the Biomedical Engineering Center (BMEC). Started in July 1999, the BMEC now has 230 researchers, most of whom were plucked from various areas within ITRI. Between them, they cover a broad spectrum of disciplines, including molecular biology, medicine, chemistry and optoelectronics.

With these multidisciplinary strengths, director Johnsee Lee believes the BMEC is uniquely qualified to pursue biochips, an



Facing change: Chintay Shih, president of ITRI.



New challenges: Taiwan's Industrial Technology Research Institute.

emerging field of biotechnology that requires materials know-how and knowledge of molecular biology. The BMEC's microfluidic biochips are different from the microarrays used in research on gene expression, Lee says. "Our chips will be for clinical application." For example, they will be used to screen for infectious diseases such as hepatitis C.

"Other institutes in Taiwan working in biotech simply do not have the expertise in miniaturization or the material sciences," says Lee. The roughly 40 biochip researchers have already produced 20 patent applications on microfluidics, probe design, surface chemistry and detection technology.

Sequential systems

Johnsee Lee also hopes to develop new algorithms to deal with the explosion of genome data. "Such high level algorithms will be helpful also as a 'fingerprint for raw materials' in sorting out the complexities of Chinese herbal medicine," he says. "We are aiming to develop one of the best sequence analysing systems in the world."

The BMEC is still a comparatively small operation by ITRI standards, accounting for only about 5% of the institute's budget of US\$540 million, and less than 5% of its researchers. Given this, and a somewhat late start, some critics think it cannot compete in DNA chips. They suggest instead that the BMEC should stick to one of its other core areas, such as medical devices and instruments and biomaterials, which can build more directly on ITRI's proven engineering and manufacturing skills.

Johnsee Lee disagrees: "While the discovery stage for DNA chips might be past, engineering, development and commercialization are still in early stages, and these are our strengths at ITRI." He plans to put together a patent portfolio and use strategic partnerships to emerge as a key niche player in the field. Other initiatives in nanomaterials and high-frequency broad-band wireless telecommunications also mark new directions as ITRI tries to redefine itself.

D.C.

ITRI ♦ <http://www.itri.org.tw>

Taiwan Semiconductor Manufacturing

Company ♦ <http://www.tsmc.com.tw>

BMEC ♦ <http://www.bmec.itri.org.tw>

where many ethnic Chinese researchers were employed, the researchers began to flood back to Hsinchu (see graph on p. 423). The park is now filled to overflowing, and a second park began operations in 1997 at Tainan in the south (see “A science giant rises out of the rice fields”, overleaf).

Some of these returnees, such as Nicky Lu, a former researcher at IBM’s Thomas J. Watson Research Center and now president of Etron Technology in the HSIP, have established start-ups back in Silicon Valley to push forward the frontiers of technology their companies are interested in (see “Nicky Lu — entrepreneur”, overleaf). Thus, a dynamic two-way interaction between Silicon Valley and Hsinchu is being established.

Academia’s rising star

The success of the electronics industry and booming Taiwan economy created an environment in which US-based Chinese scientists in basic research also began to return.

One example is Nobel prizewinner Yuan-tseh Lee, who came back to head the Academia Sinica, Taiwan’s leading institution for basic research, in 1994. In turn, Lee

has been instrumental in bringing back the brightest and best to the academy, which is now beginning to make its mark in molecular and structural biology, astronomy, biomedical science, surface science and Lee’s own field of molecular and atomic science (see “Beating the odds”, p. 422).

The academy, which consists of 25 institutes housing 818 researchers, has increased the number of its papers published in international peer-reviewed journals from 200 to 1,220 in the past decade. This gives a 14% share of the total papers in the Science Citation Index that Taiwan produced in 1998.

As well as strengthening the academy by implementing recruitment campaigns and external reviews of staff and institutes, Lee has established a technology transfer office. For the first time, scientists are seeing their ideas patented and developed by local industry.

In the first 70 years of its existence, the academy failed to secure a single patent. But since 1998 it has filed 80 applications in the United States, of which ten have been approved and seven have been licensed to companies. The technology transfer office already has about US\$7 million to show for its

efforts, moving into the black after just three years. The office also helps with the licensing of research and holds exhibitions to showcase Academia Sinica research to potential buyers.

Microarrays or ‘biochips’ are one example of success. Konan Peck of the Academia Sinica’s Institute of Biomedical Sciences has licensed his colorimetric detection technique for complementary DNA microarrays to three local start-ups — U-vision Biotech, Taiwan Genome Sciences and GeneMaster Biotechnology Corporation — and one US start-up, PhytoCeutica, which was established by Yung-chi Cheng at Yale University.

But the academy and its scientists are new to the game of technology transfer, and not accustomed to lengthy patenting processes, which can severely delay publication of research papers. One of the academy’s most commercially promising results — a mutant mouse for studying spinal muscular atrophy — illustrates some of the frustrations researchers unused to the business world face when trying to capitalize on their research (see “The rodent route to celebrity”, p. 425).

Overall, technology transfer remains underdeveloped. Chi-ming Liang, director

Building a progressive sci-tech island

In 1949, a rag-tag Chinese nationalist army (Kuomintang or KMT) fled the communist army of Mao Tse-tung and took refuge on the island of Taiwan — officially taking the name Republic of China. The KMT controlled the island until the election of Chen Shui-bian of the Democratic Progressive Party (DPP) in March this year. The mainland People’s Republic of China maintains a “one China” principle, stating that Taiwan is just a renegade province. Although relations between the island and the mainland have always been tense, the election of a member of the DPP, which is known for making strong statements favouring Taiwanese independence, has irked mainland leaders.

The influx of ethnic Chinese from abroad, mostly but not exclusively returnees of Taiwan origin, has fuelled Taiwan’s electronics boom, and many of the highest rungs of Taiwan’s academia are populated by Chinese who have come, or returned, to Taiwan in the past decade. The government has dedicated itself to making Taiwan a “progressive sci-tech island”. It has already shown ingenuity in establishing science parks and industrial institutes, which gave rise to the successful electronics industry. The government budget for science and technology for 2001 stands at NT\$53.1 billion (US\$1.7 billion), up 10.3% on 2000.

The government continues to promote both academic and industrial R&D. Public R&D spending has increased yearly by 8–10% for each of the past five years, and aggressive government policies have encouraged private research so that total R&D spending (government plus industry) has jumped from 1.0% to 2.0% of GDP since 1986. This puts Taiwan on par with the United Kingdom, France and Germany but still behind the United States, Japan and Korea. This has boosted academic science, though some see an “overly democratic” emphasis on even distribution of research funds as preventing development of some of the best talent. But Taiwan’s international openness has created a fluidity in terms of capital, technological and human resources that has made it a vital node in the global economic, technological and scientific web.

D.C.



A science giant rises out of the rice fields

Hot on the heels of Taiwan's hugely successful science and industrial park in Hsinchu comes the Tainan Science-based Industrial Park (TSIP). Starting operations in 1997, the TSIP is designed to develop and upgrade high-tech industry in the south of the island.

It has a tough act to follow. With government funding for infrastructure and administration, and a combination of tax-breaks and subsidies for resident companies and plants, Hsinchu is now a massive 292-company park with an annual trade value of US\$21.76 billion.

By contrast, the TSIP's location is remarkably rural, with its shiny new buildings rising above a vista of rice and sugar fields. In fact, Tainan was chosen partly because it was close to various agricultural research institutes such as the Asian Vegetable Research and Development Center, Taiwan Livestock Research Institute and the Tung Kang Marine Laboratory. The park's planners originally hoped that agricultural biotechnology would be the main discipline in evidence at the TSIP.

Sowing the seeds

But the reality is somewhat different. "We were trying to take advantage of the research institutes and agriculture-based industries that were already well-developed in the south," says Chein Tai, the TSIP's director general. "But of those agrobiotechnology companies that applied for park space, none had a high enough level of technology."

"We tried to support agrobiotechnology, but we can't do more than give infrastructure," explains Steve Hsieh, vice-chairman of the National Science Council, which administers the science parks. "The core products at Tainan are electronics and information technology, thin-film transistors and liquid crystal displays."

In fact, Tainan's industrial composition now looks

suspiciously similar to its northern counterpart in Hsinchu, with semiconductor foundry giants Taiwan Semiconductor Manufacturing Company and United Microelectronics Corporation among the first to set up new plants. Of the 40 companies currently moving to Tainan, only four are involved in biotechnology, the rest are split across integrated circuit design and manufacturing (16), computers and peripherals makers (2), wireless and telecommunications (5), optoelectronics (11) and precision instruments (2).

Growth factor

The National Cheng Kung University (NCKU), one of Taiwan's top four universities, is close to Tainan. It hopes to play a similar role at the TSIP as Hsinchu's National Tsinghua and National Chiao-Tung universities do at Hsinchu, by fuelling the park's growth with well-trained workers, technology transfer opportunities, and an incubator programme.

Given that Tai was director of the NCKU's biomedical research centre, close collaboration can be expected. The NCKU is also relatively progressive in its attitude towards letting academics collaborate with industry. The university allows professors to work with start-up companies in one of two ways — they can work with the company for eight hours per week while carrying on university teaching and research responsibilities, or they can take a sabbatical of up to three years while the university holds their position open.

But as yet there are few tangible examples of interaction. Jyh-ming Ting, of the NCKU's Micro-Electro Mechanical Systems centre, hopes his technology will take root in a TSIP start-up. Among other things, Ting works on chips for high-throughput analysis of protein interactions. But he admits that he shares



Rural revolution: Tainan's science park is fuelling high-tech industry.

TSIP

some of the conservative attitude that has kept other NCKU academics from cashing in on the university's lenient industrial collaboration policy. "No one is willing to jump in the water first," he says.

The TSIP has other advantages as well as its location. For a start, it is spacious. Even at full capacity, the park has been designed to avoid the overcrowding seen at Hsinchu. Tainan also offers lower living costs and a cheaper, although still high quality, labour force compared with Hsinchu, says Hardy Chan, of ScinoPharm Taiwan, a company producing custom-made ingredients for pharmaceuticals. "There are many top quality technical schools in middle and southern Taiwan. In the past many graduates had to go north to find work, but they are happy to stay in the south," he says.

Biotech success?

ScinoPharm took advantage of the government's increased interest in biotechnology to build expansive facilities at the park. Launched in 1994, ScinoPharm released its first product only this year, and some investors are getting anxious.

But Jo Shen, president of

ScinoPharm and a veteran of the US pharmaceutical industry, is not concerned. "Many Taiwanese investors try to think of biotech in terms of the quick return possible in semiconductor manufacturing, but we have to operate under a different model. You need facilities that will make the customer feel confident." Indeed, many researchers feel that investor demand for short-term return is a stumbling block for the development of biotech in Taiwan, and ScinoPharm (already with a backlog of orders) is one company to watch to see if the long-term vision can prove itself in Tainan.

The pace at Tainan seems set to pick up soon, with the Taiwan Semiconductor Manufacturing Company moving its corporate headquarters there and the Industrial Technology Research Institute also setting up shop.

All of Tainan's lots are accounted for, and Tai is already predicting a labour squeeze by 2004. But by then, he forecasts that the TSIP will have 100 companies and a production value of US\$10 billion — well on its way to being another Hsinchu.

D.C.

► <http://www.sipa.gov.tw>

NCKU ► <http://www.ncku.edu.tw>

of the office, is a case in point. His post is merely a moonlighting job to his position in the academy's Institute of Biochemistry. It earns him only a paltry salary supplement despite the heavy load in terms of time commitment and responsibility.

To compound the problem, some researchers still object to the sale of their ideas and are willing to give them away. Liang recalls one researcher who off-handedly offered to license his technology for NT\$1, sending Liang scrambling to renegotiate a sum that would at least cover his office and patenting expenses.

Taiwan also lacks potential leaders with multidisciplinary backgrounds combining scientific, legal and business knowledge. Liang prefers to entrust most legal matters to lawyers in the United States. He gets patents written in the United States because, he says, many Taiwanese lawyers will "write very narrow patents, US lawyers know how to make the broadest possible coverage".

Universities galore

Ironically, just as the academy tries to move closer to industry the very success of Taiwan's industry and the growth of the Taiwan economy are conspiring to make life difficult for basic research. With the boom of the economy over the past decade (except for a short downturn in the Asian economic crisis), many new universities have been established in Taiwan, most of them upgraded institutes and colleges.

"There are so many universities I am losing count," says Steve Hsieh, vice-chairman of the National Science Council (NSC), Taiwan's government organization that provides competitive research grants to universities, the Academia Sinica and other government research institutions. As a result, the NSC has had to spread its funds over an ever widening base. The average size of grants has actually declined during the 1990s, despite a rapidly rising grant budget for the NSC.

The council is trying to counter this with its new frontier grants, targeted at only the very best researchers in Taiwan. But more still needs to be done to separate the many universities doing relatively mundane research from the few truly outstanding ones.

Some top academic advisors to the government, for example, advocate research assessments similar to those in Britain's university system, which would help target a larger share of block funding at only a small number of top institutions doing outstanding research. Understandably, these ideas are meeting strong resistance from most of Taiwan's 42 national and private universities. But they are welcomed at those institutes, such as the NTU, that are leading the field.

Another consequence of the university boom is that many Taiwanese students now choose to stay in Taiwan rather than go overseas. About 90% of students now stay in Taiwan for graduate studies or to join industry, says Hsieh. And the make-up of those going

overseas has changed markedly, with a growing number favouring business studies over science and engineering.

Cheng-ching Li, director general of the bureau of international and cultural relations at the Ministry of Education, is concerned about this shift. Taiwan has benefited greatly from the return of former overseas students who have gained broad experience in the West and have fluent English language skills. The ministry is now offering scholarships to encourage the best students to go overseas. But it can support only a few hundred students.

The domestically trained students have a very different attitude to those who have spent time overseas. "Most do not want to do PhDs," says Kopin Liu of Academia Sinica's Institute of Atomic and Molecular Sciences, which is located on the NTU's campus. "Most stop at the Masters level and go to industry."

Even those who take up postdoctoral positions often leave for industry after only one year of research, says Liu. This causes discontinuity in the research programmes at the institute. "On one occasion we lost five postdocs overnight to an IT start-up in Hsinchu, or 10% of our whole complement of 50 postdocs." Recruiting replacements is not easy.

Most of those who leave for industry do not move to research positions but go into areas such as marketing and sales. "The salary in industry is not much better. It is the stock options and bonuses that attract them," says Liu, explaining the departure of one of his postdocs to sell vacuum pumps.

Adding to these local problems, the recent upturn in the US economy has reduced the number of ethnic Chinese prepared to return to Taiwan. Rising political tensions with the Chinese mainland have also deterred potential returnees. The overseas supply is drying

Nicky Lu — entrepreneur

It started with a straightforward offer. "You give me US\$10 million over five years to set up a company and I get to innovate — to make new chips and technology."

Under these simple terms, Nicky Lu established Etron Technology, a company that has taken Taiwan to the cutting edge in the design of semiconductor memory chips. His innovations even won Lu the 'Nobel prize' of electrical engineering in 1998 — the Solid-State Circuits Award from the Institute of Electrical and Electronics Engineers.

Born in Hong Kong, Lu graduated from the National Taiwan University in 1975 and obtained a PhD in electrical engineering from Stanford University. He made a name for himself at IBM's Thomas J. Watson Research Center in New York by tripling the access speeds of DRAM (dynamic random access memory) chips from 60 nanoseconds to just 20 nanoseconds.

This brought Lu to the attention of Taiwan's Industrial Technology Research Institute, which asked him to join



its submicron project aimed at developing expertise in DRAM chip technology. It was then that Lu, imbued with an entrepreneurial spirit learned at Stanford, set out the terms for his new business.

Etron is now an industry leader in memory products. It is the main supplier of DRAM chips to Intel, maker of the Pentium series processors found in most of today's desktop and laptop computers. Lu continues to position his business interests at the forefront of the electronics industry. He has established offshoot companies in both Taiwan and Silicon Valley, devoted to

cutting-edge fields such as wireless communications and semiconductor testing, the latter of which Lu describes as "the real technological bottleneck in semiconductor production".

Lu's business success belies his respect for research for its own sake. He recognizes that basic research can lead to technological strength. "Unlike the United States, Taiwan lacks people dedicated to the search for excellence, regardless of money, or fame," Lu says. "Taiwan needs more 'nerds' dedicated to innovation."

D.C.

Etron ♦ <http://www.etrn.com>

regional insight taiwan

up admits Yuan-tseh Lee, who has had difficulty finding a head for a new centre of applied science and engineering at the Academia Sinica. Competition with local industry has made recruitment for this particular position unusually difficult.

But Lee remains confident about the future. Some top Chinese scientists from the West are still willing to join the academy, and Lee is expecting to appoint a new head of the Institute of Biomedical Sciences shortly.

The biotech future

Lee sees particular promise in genomics, where the Academia Sinica has several rising stars. The academy is planning to establish a special centre for functional genomics which will be open to several of the academy's institutes, and Lee says that five investors from the microelectronics industry are prepared to put US\$100 million into the centre over the next five years.

According to officials promoting biotechnology, there are billions of dollars in the private sector waiting to find a home in biotechnology. But will Taiwan be able to put the funds to productive use?

Just as with electronics, the Taiwanese government has tried to develop biotechnology over the past two decades. But the results so far stand in stark contrast to the booming

electronics industry. One only needs to look at Hsinchu to see the tiny contribution biotechnology makes — 0.1% of sales (see Table 1) — and so far there is no sign of a major upswing.

The Development Center for Biotechnology (DCB) was established by the government in 1984 with a mandate to help develop and nurture a biotech industry. After 16 years and hundreds of millions of dollars poured into hundreds of small projects, there is nothing to show. The centre's annual budget has been frozen at the same level of about US\$16 million for several years following severe criticisms in government and academic circles. A new director has just been appointed to rectify the situation, and in a show of confidence the government will boost the DCB's budget next year by 40 per cent to US\$23 million.

The new DCB president, Tse-wen Chang, former dean of life sciences at National Tsinghua University near the HSIP, has a proven track record. He recently returned to Tsinghua from the United States where he and his former wife set up Tanox, a successful company that co-developed an asthma drug called Xolair with Genentech and Novartis. The drug has completed phase III clinical trials and is expected to get marketing approval from the US Food and Drug Administration (FDA) soon.



Tse-wen Chang: guiding biotech investment.

Chang plans to phase out the hundreds of discovery projects taken on by the DCB over the years. In a complete redefinition of the centre's mission, he will spend the next two years looking for 15–20 promising technologies in the United States and elsewhere in the West that are suitable for joint development with industry in Taiwan. Pharmaceuticals will be a key focus.

To help this process, the DCB is setting up five liaison offices — four spanning the US east and west coasts and one in Cambridge in the United Kingdom — manned by Chinese scientists who will look out for promising technologies being developed by start-up companies. The DCB will then seek partners in local industry in Taiwan to co-develop the products with the western start-ups.

Beating the odds

Nobel prizewinner Yuan-tseh Lee's greatest achievement since taking over at the Academia Sinica, Taiwan's premier academic research institute, in 1994 has been his recruitment of a force of world-class Chinese scholars.

Among these are "10 or 15 key people — not just scholars, but leaders", according to Sunney Chan, the academy's vice-president. "Lee is developing a culture." Even those of Chinese descent without ties to Taiwan, such as Chan, a native Californian, and the academy's Institute of Astronomy and Astrophysics (ASIAA) director Fred Lo, who was born in Hong Kong, are casting their lots with Lee.

Lee's success is all the more startling considering the impossibility of competing with salaries in the United States. All university and Academia Sinica researchers are considered government employees and so get a fixed salary.

"Universities and research

institutes in the United States have much more flexibility to give extra to top scholars," explains Yuen-Ron Shen, a physics professor at the University of California at Berkeley.

Kopin Liu, for example, had to swallow a 66% salary cut when he moved to the academy's Institute for Atomic and Molecular Science from the Argonne National Laboratory in Illinois. Liu says he came because he could both do research and help educate Taiwan's younger generation. But he still had mixed feelings: "It's hard to tell your wife and kids, 'I have a dream to follow, but you must make sacrifices with me'," he says.

Salary shortfalls

Lee took a first stab at the salary problem in 1994 by establishing the Foundation for the Advancement of Outstanding Scholarship. Local private companies have contributed



All change: the Academia Sinica is redefining itself.

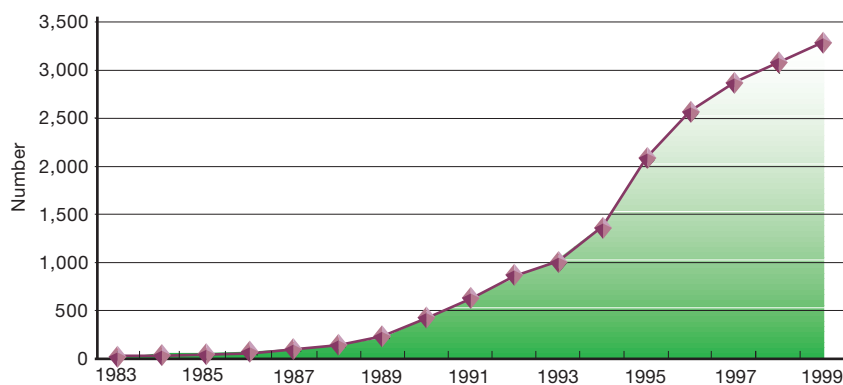
NT\$1 billion (US\$32 million), enough for a self-supporting endowment. The foundation offers renewable five-year grants to researchers and administrators throughout Taiwan's research institutions. At full capacity it will have about 100 funded scholars.

The money is aimed specifically at bringing earnings more in line with salaries abroad. Though they are grants, for many scholars they are considered a necessary salary

supplement. Liu thinks the five-year term is a good idea. It will "prevent people from coming back to Taiwan for an easy retirement", he says.

Profitable work

Another perk now being offered is the possibility of capitalizing on one's research, which until last year was against the law in Taiwan. Lee worked for the introduction of a law allowing researchers in Taiwan to profit from their results, and he established a technology



Welcome return: the number of ethnic Chinese researchers coming to Hsinchu from overseas.

“Investors are abundant in Taiwan but they don’t know what technology to invest in,” says Chang. It will be the DCB’s role to act as a bridge in introducing them to promising technologies and help with their development, he explains.

Ming-chu Hsu, director of the Division of Biotechnology and Pharmaceutical Research at Taiwan’s National Health Research Institutes (NHRI), is adopting a similar approach (see “Homeless, but heading for success”, p. 424). She believes that Taiwan has the potential to be a key centre for the clinical trial of drugs

aimed at the potentially huge Asian market.

There are several diseases prevalent in Asia, such as liver and cervical cancer, which receive limited attention from western pharmaceutical companies. Hsu sees great potential for Taiwan in helping to develop therapeutics and diagnostics for these conditions.

Taiwan was the first country in Asia to adopt US FDA rules for clinical trials and has established a Center for Drug Evaluation with the help of FDA staff. Hsu also points out that Taiwan is unique in that the Ministry of Economic Affairs can provide substantial

support for clinical trials. Half of the costs of phase I trials can be covered by the government, and the state will even pay 25% in grants and 25% in interest-free loans for the final phase III stage of clinical trials where the drug is tested out on large populations of patients.

The ministry has set aside US\$650 million for investment in the development of biotechnology over five years from 1998, focusing mainly on pharmaceuticals and agrobiotechnology. Hsu says the private sector is willing to match each government dollar with two or three dollars, bringing the total to about US\$2 billion. Chang estimates that the total investment figure for biotechnology over the next five years (from 2000) could be even higher at between US\$5 billion and US\$6 billion.

But although Taiwan has gained experience in phase III trials by working with western pharmaceutical companies, “we need practice at phase I,” says Hsu.

Her strategy in the short term is to identify drug candidates, technologies or products for licensing to Taiwan and thereby shorten the time to market. By doing this, Hsu plans to create an infrastructure for drug development, and so move on to developing her own drugs. She recently visited five US biotech companies to explore opportunities for licensing their technology for development in Taiwan.

transfer office at the Academia Sinica to aid the process.

Lee also promotes the ability to interact in English, which he considers increasingly important for scientific achievement. But his proposal to establish an international graduate school at the Academia Sinica led to a showdown with the Ministry of Education and many top universities, which maintained their sole right to confer degrees.

A compromise has been reached for a new collaborative graduate programme in which the entrance exams and course are in English. The programme will be administered and taught jointly by the Academia Sinica and selected universities, but the latter will give the degrees.

Many academy researchers hold joint appointments in a university and will naturally be instrumental in bridging the gap between the academy and the universities. K. Y. Lo is one such joint-appointee with a position

at the National Taiwan University’s department of physics and the academy’s ASIAA, but he is also a natural for another reason: his institute already conducts all of its meetings and discussions in English.

The collaboration key

Lo’s institute promotes a different kind of internationalization in its collaborations on telescopes. Work with the Smithsonian Institution on the Submillimeter-wave Interferometric Array in Hawaii has brought the ASIAA international recognition.

Early this year, with Ministry of Education funding and in conjunction with the National Taiwan University, Lo started the Array for Microwave Background Anisotropy project, a collaboration with the United States, Canada and Australia. Lo lists a number of areas where he believes the project can make an impact, such as

examining the problem of missing baryons.

Lo also sees great potential in collaboration with mainland China, and this spring he signed an agreement to work with Sheng-cai Shi and Ji Yang of the Purple Mountain Observatory on Tianbao Peak in China. “There are many areas where our expertise could complement each other,” Lo says.

Lee is keen to introduce more joint projects that reach outside the island, but he also feels it is important to promote collaboration within the Academia Sinica. A number of centres for joint use by academy institutes in areas such as agrobiotechnology, functional genomics and applied science and engineering are being created to encourage these emerging multidisciplinary fields.

But Lee has encountered not only the walls of a traditionally isolationist academic community but also

laws that prohibit such centres. Many advocates of the centres, however, seem to have been successful in making their claim to the government, and the centres are expected to become legal this month.

The centres will likely spur new research and, with the help of the technology transfer office, more interaction with industry. But they also pose administrative problems for the academy. For example, in assessing member’s performances, academy officials will have to figure out how to evaluate members who have patents pending and so have not yet published their work.

The Academia Sinica is in a state of rapid change. There are still problems, foremost among them a chronic shortage of postdoctoral students. But Lee, in a steady methodical fashion, seems to be handling these difficulties one by one.

D.C.

♦ <http://www.sinica.edu.tw/index.shtml>

Homeless, but heading for success

It may be homeless, but Taiwan's National Health Research Institutes (NHRI) is having a profound effect on medical research on the island. In its seven-year existence it has encouraged networking among Taiwanese scientists and stimulated other government organizations to review the way they fund research.

The NHRI has its roots in a competitive extramural grant programme set up just over seven years ago to enhance the standards of medical research throughout Taiwan. The programme has achieved a modest scale of about NT\$300 million (US\$10 million) a year that belies the extent of its impact.

Cooperative movement

The NHRI has been active in networking with top medical organizations throughout Taiwan to help raise standards of research and clinical practice. It has clinical research wards with research beds and associated laboratories at Taipei Veteran General Hospital and the National Taiwan University Hospital. It has also started the Taiwan Cooperative Oncology Group, a multi-institutional network for cancer clinical trials in which the Institute of Biomedical Sciences and nine teaching hospitals throughout Taiwan are participating.

NHRI grants have set new standards in Taiwan. They are peer-reviewed internationally and have a high rejection rate of 70–80%. Around 40 grants are awarded annually, each worth US\$100,000–300,000 per year for up to five years.

By contrast, the National Science Council accepts over 70% of the roughly 15,000 grant applications it receives. As a result, it can only supply piecemeal, short-term awards that are about a tenth the size of an NHRI grant. The council has now introduced frontier research programme grants, to provide more substantial, long-term

support to outstanding work.

The most ambitious component of the NHRI is its intramural programme. This has grown to 460 personnel including affiliated institutions and has an annual budget of about US\$32 million. Yet the organization still does not have a place it can call home.

The lodger

Since 1995, the NHRI has been “temporarily” housed on three floors of the Institute of Biomedical Sciences at the Academia Sinica, where Cheng-wen Wu, NHRI president, used to be head. Many personnel are dispersed at six other locations.

Wu has inspected various plots of land for the NHRI over the past few years, but it was only early this year that he settled on a 32-hectare plot south of Taipei, near the Hsinchu Science-based Industrial Park. Wu hopes to move the NHRI there by 2003 and house up to 1,000 staff in three research buildings in the first phase of development.

Some top government advisors fear the location is too isolated, but others see potential for the NHRI to interact with universities in the Hsinchu area. In particular it could forge links with the National Tsinghua University, which has ambitions to strengthen its medical research but lacks a research hospital.

New directions

Wu has found directors for eight of the ten divisions planned for the NHRI: clinical research; biotechnology and pharmaceutical research; biostatistics; cancer research; molecular and genomic medicine; health policy research; environmental health and occupational disease; and biomedical engineering. But he is still searching for heads of gerontology, and mental health and drug abuse. Some government officials have criticized Wu for taking so long to fill these posts. Although Wu says he is determined to wait

until he can find good leaders.

But recruitment has become more difficult in recent years. Fewer students in Taiwan are going overseas to gain valuable experience in western research institutions and fewer ethnic Chinese scientists in the West want to take up jobs in Taiwan because of the political instability in relations with the Chinese mainland. “They were enthusiastic ten years ago about Taiwan, but now people are worrying about the continuous threat from China,” says Wu.

Working together

Nevertheless, Wu has managed to recruit some top leaders, such as Ming-chu Hsu, a former research director of oncology at Hoffmann-La Roche, who now not only heads the Division of Biotechnology and Pharmaceutical Research but also leads a national programme for those disciplines.

She reports no difficulty in recruiting staff for her division, with applications from over 400 PhDs from the United States, Germany, Eastern Europe and China. But she agrees that the Taiwan still needs experienced leaders.

A key aim of the NHRI is to draw together the limited numbers of researchers in Taiwan, who tend to work in isolated groups. “There are only 3,000 to 4,000 biomedical researchers in the whole of Taiwan, which is about the same as in one large western pharmaceutical company,” says

Wu. So it is essential for them to join forces if they are going to have an impact.

Hsu reports success on this front. In just one year, she has managed to bring together 40 synthetic chemistry and natural products laboratories throughout Taiwan to create a chemical library that will be crucial in the search for novel drugs. “It wasn’t so difficult as some people suggest. We just had to go and talk to them and bring them together,” she says.

Similarly, during an enterovirus pandemic that caused more than a million cases of hand, foot and mouth disease in Taiwan in 1998, the NHRI and three universities pulled together to form a core laboratory that isolated the virus responsible and developed a vaccine. The results were reported in the *New England Journal of Medicine* (341, 929–935; 1999).

But the road to success has not been easy for Wu. In one of his most exasperating experiences, he was sued after firing a staff member found to be lacking in a review. It took three years to settle the case in Taiwan, where such removal of an academic was unprecedented. Despite this, the NHRI is beginning to have a significant effect, and once it has a permanent home it will be ready to push medical research, clinical practice and pharmaceutical research to greater heights in Taiwan. **D.S.**

NHRI ♦ <http://www.nhri.org.tw>

Taiwan Cooperative Oncology Group ♦ <http://www.nhri.org.tw/cancer/en/tcog.htm>



Ming-chu Hsu: bringing researchers together for drug research.

Jen Chen, general manager of Genelabs Biotechnology, a start-up company in Hsinchu that is performing clinical trials for western pharmaceutical companies, agrees that clinical trials for Asia are a niche that Taiwan can develop in biotechnology. But he points out that the severe shortage of doctors who know how to do clinical trials is a serious problem for Taiwan. Genelabs employs its own doctor trained in clinical trial practice to lead projects at hospitals.

"Only when the presidents of top medical institutions, such as the medical school of NTU and Taipei Veterans General Hospital, stand up and say that they are going to focus on clinical trials will things begin to happen," says Chen.

Others are more sceptical about the development of biotechnology in Taiwan. John Yu, first vice-president and general manager of the technology department of the China Development Industrial Bank, the largest investment bank in the Asian-Pacific region outside of Japan, agrees that private funds for investment in biotechnology have grown significantly in recent years. Two-fifths of the US\$50 million investment fund Yu oversees is earmarked for biotechnology. The fund has risen to this level rapidly over the past three years, with most activity taking place this year.

But all of Yu's major biotech investments are in the United States and mainland China, and he sees little opportunity for worthwhile

local investment at present. "People neglect the fact that the semiconductor industry and biotech are totally different. We don't have the brain power or environment to develop biotechnology," he says.

Switching sides

In general, academics in Taiwan are unwilling to get involved in venture businesses, and the few who do take the plunge find themselves hampered by government regulations. The rules have been eased in recent years, allowing academics to share in the spoils of intellectual property rights. But the ability to work in industry, which depends on respective university's policies, is severely restricted with some of the more lenient universities allowing a sabbatical or a mere 8 hours per week to devote to an industry position. For those serious about their enterprise, the solution is often to leave academia all together.

Rong-hwa Lin, for example, recently resigned from his job as director of the Graduate Institute of Immunology at the National Taiwan University College of Medicine to defect to the private sector. Together with Herbert Wu, an associate professor in the Institute of Molecular Medicine, he is putting the finishing touches to their start-up company, AbGenomics.

Lin and Wu have devised a method for preparing antibodies using DNA immunization. Although many researchers around the

world have dismissed this type of technique, AbGenomics is already establishing collaborative agreements with big-name research institutes and universities in Taiwan and in the West, including with the National Institutes of Health (NIH) in the United States.

The NIH showed interest after Lin and Wu generated antibodies to two cancer-related proteins from genes it supplied. The NIH had been trying in vain for three and a half years to prepare the antibodies; Lin and Wu succeeded in less than two months.

But Lin remains concerned about the university regulations he feels impeded his research, especially the government stipulation that after three years a research assistant cannot receive another pay rise. With such poor prospects, few stay for long. "I felt like I was spending half my time retraining research assistants," Lin says. In the private sector, Lin and Wu can pay research assistants according to their performance, rather than just a standard low wage. "Research assistants can treat it as a real job — one they'll want to stay at," Lin says.

Fleeing the regulations of academia, Lin and Wu were also wary of restrictions often imposed by investors. They chose investors who understood biotech and who did not blink when told that they might be looking at a 3–5 year start-up period. The fact that they could be selective points to the startling amount of money available for biotech.

The rodent route to celebrity

Mice have made Hung Li a celebrity in Taiwan's science community. In January this year, Li's paper detailing his mouse models for a fatal motor neuron disease appeared in *Nature Genetics* (24, 66–70; 2000). On a small island where scientists have only recently started publishing in major journals, this caused quite a stir.

But drowning out the domestic acclaim has been an inundation of requests from overseas laboratories seeking

the licensing rights to Li's mice and his results.

Li's research, carried out at his lab in the Academia Sinica's Institute of Molecular Biology, has specific applications for studying spinal muscular atrophy (SMA). Usually appearing in infants, this disease weakens the muscles and often leads to respiratory-related fatality.

To tackle the problem, Li generated genetically engineered mice that suffered

from a form of human SMA. He then used them to screen drug candidates. Two patents are now in the offing, and biotech companies in the United States and Taiwan are offering around US\$1 million for the mice and/or related research results.

But Li would have to share any royalties with his department, the academy and the government. More importantly, any negotiations must be done through the academy, and Li must follow their stipulations concerning publication of his research.

With the patent for the mouse model safely filed, it is now available to drug screeners, and the Academia Sinica has recently signed a contract allowing a Taiwanese biotech start-up to use the mice. But the SMA drug patent is still sitting in a lawyer's office in the United States, waiting to be filed.

Li is learning an unpleasant

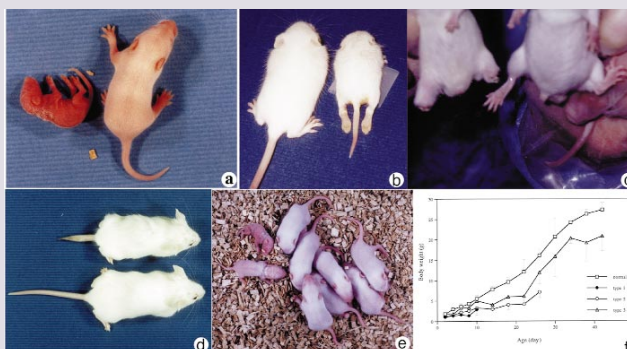


Hung Li (second from right) and the team behind the mice.

lesson about the effect of patent red-tape not only on licensing but also on research publication. Because Japan and the European Patent Office do not recognize a patent filed after the research underlying it has been published (unlike the United States, which offers a one-year period to file), Li must wait to publish his drug research data.

Li is not so concerned about the money. But he is somewhat vexed by the delay caused by the patenting process and the subsequent accusations made by biotech companies that he is holding up drug development.

D.C.



Million dollar mice: Li's rodent models have huge potential.

Indeed, as they gathered the US\$16 million to start AbGenomics, Lin and Wu turned down about half the capital put up by investors.

Taiwan also lacks lawyers with training in biotechnology issues, and venture businesses have to seek the help of US lawyers. So it is not surprising that many investors in Taiwan look to the United States to place their funds, and this is where the network of Chinese scientists in the United States comes into play.

The China connection

The biggest current investment made by John Yu's department is US\$2 million in the start-up company Aviva Biosciences. Based on the west coast of the United States, the company is developing biochips designed by Jing Cheng of Tsinghua University in mainland China (see "Bridging the gap", right).

The bank is also planning to invest in several 'incubators' to develop new technology in mainland Chinese cities under the umbrella name Chinalab. The first is in Beijing and covers software development, a second will soon open in Shanghai. In all, ten incubators in ten cities are planned with each incubator matched to the high-tech strength of the city. The bank has also invested in Yung-chi Cheng's company, PhytoCeutica.

Yu's belief is that a network of ethnic Chinese scientists linking the United States, the Chinese mainland and Taiwan could be very powerful in the development, manufacture and marketing of biotechnology products.

Yuan-tseh Lee is also active in trying to develop links with the Chinese mainland. There are about 50 mainland Chinese scientists currently at the Academia Sinica. But under the strict regulations imposed by the mainland government, the visitors can only stay a maximum of one year. Also, because Taiwan is considered to be a renegade province by the mainland, the researchers are not granted a passport, preventing them from travelling to scientific meetings outside of Taiwan. This severely hampers their ability to have an impact during their short stays.

Two years ago, Zhu Lilan, China's minister of science and technology, became the first ever minister from the mainland to visit Taiwan. This trip was taken as a sign that thawing relations between Taiwan and the mainland might lead to greater scientific exchange.

But the election of pro-independence president Chen Shui-bian in Taiwan, who received crucial election backing from Lee (see *Nature* 404, 917; 2000), has brought a new chill to relations across the Taiwan Straits. Thus, direct exchange and interaction in science and technology between Taiwan and the mainland seems likely to remain limited for the foreseeable future.

For much of the post-war period, Taiwan focused on building its industry, improving education and defending itself from possible

Bridging the gap

Political tensions might be high between Taiwan and mainland China, but that has not stopped some enterprising scientists from bridging the divide. One such entrepreneur is Jing Cheng, director of the new biochip research and development centre at Tsinghua University in Beijing.

Cheng is developing 'active' electromagnetic biochips that can be used in bioassays. They are much faster and more sensitive than the 'passive' chips such as glass microarrays, he claims. In addition, Cheng says his biochips can use traditional salt buffers, giving them an advantage over their electronic counterparts

which need special low-salt buffers.

Cheng was working at Nanogen, a US biochip venture company in San Diego, when he was lured back to Tsinghua to set up and head the biochip centre with more than US\$4 million in support from the university and grants. But even as he moved back to China, Cheng was raising funds to set up a venture business on the west coast of the United States to develop his biochips (see *Nature* 399, 178; 1999).

The biggest investor in Aviva Biosciences, Cheng's new San Diego-based start-up, is the China Development Industrial Bank in Taiwan. The bank

has contributed US\$2 million of the \$5 million raised to launch the company and is represented on its board.

To get his chips on the market, Cheng has forged an alliance between mainland China, Taiwan and the United States. Under the plan, Cheng's group at Tsinghua will progress discovery research, Aviva will develop the chips for market, and the chips will be made in Taiwan. The chips might also be manufactured on the Chinese mainland "at a later stage", says Cheng. Tsinghua University will retain the rights for the licensed patents in the mainland, while Aviva will hold all the rights elsewhere.

Aviva president Julian Yuan, who comes from Taiwan, expects the first chips to be ready for market in about three years. Cheng says they are already doing their homework to find a suitable pharmaceutical or diagnostic company to form a partnership with when the chips are ready. **D.S.**

► <http://www.tsinghua.edu.cn>



Working together: Tsinghua University is at the heart of a joint venture between China, Taiwan and the United States.

attack by the Chinese mainland. Now it is reaping the rewards from its strategy. The gains made in electronics have put Taiwan firmly on the world map, but its ambitions in biotechnology are still quite some way from being realized.

There are many fundamental weaknesses and deficiencies in Taiwan's infrastructure that must be addressed before there can be any hope of a significant biotechnology industry. The number of life scientists in the whole of Taiwan is about the same as in one western pharmaceutical company. The obvious solution is to get more but, until that is possible, Taiwan will have to be very selective about which projects it takes on.

Taiwan also lacks the powerful pharmaceutical companies of the West, so it will have to work with them while building its own infrastructure. The leaders trying to push biotechnology forward are well aware of

these problems, and as the age of genomics dawns, the capacity of Taiwan to rise to the challenge and find a niche in this new market should not be underestimated. ■

David Swinbanks is Asia-Pacific Editor of *Nature* and David Cyranoski is Asian-Pacific Correspondent.

Web links

ITRI ► <http://www.itri.org.tw/eng/index.html>
 Science-based Industrial Parks ► <http://www.sipa.gov.tw>
 Academia Sinica ► <http://www.sinica.edu.tw/index.shtml>
 NHRI ► <http://www.nhri.org.tw>
 DCB ► http://www.dcb.org.tw/e_index1.html
 National Science Council ► <http://www.nsc.gov.tw>
 China Development Industrial Bank ► <http://www.iaib.org/china.htm>
 National Taiwan University ► <http://www.ntu.edu.tw/English>
 National Tsinghua University ► <http://www.edu.tw/bicer/english/tsinghua.htm>
 Science and Technology Information Center ► http://www.stic.gov.tw/eng_home.htm

Raw deal?

Poor exchange rate hampers European mobility for postdocs

p427

Career opportunities

Spanish postdocs are finding it hard to get permanent positions

p428

New directions

Many US postdocs are turning their backs on academic research

p429

Unfavourable economics put postdocs across Europe under strain

Travel may broaden the mind, but for many postdocs taking fellowships in the European Union, it is causing financial hardship, says Natasha Loder.

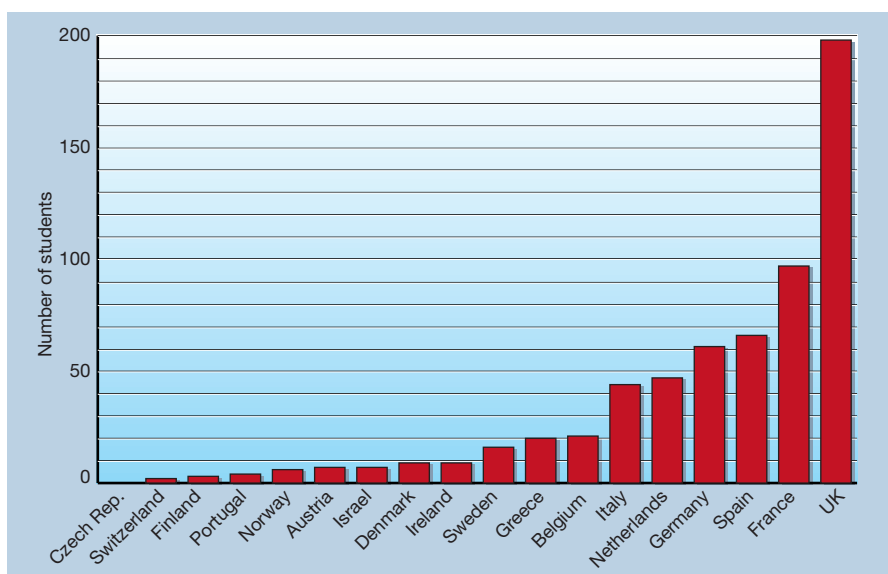
Would you accept a fellowship where currency fluctuations decreased the value of your grant by 30% by the time you reached the end of your research? Or if you were forced to pay money into a worthless pension plan? Or if you had to wait three months to be paid anything and a year before being paid in full? Then again, perhaps you would like 75% of your travel allowance to be taxed into thin air because nobody knows how to calculate what you are actually owed?

These are the realities of being a truly European young researcher in the year 2000 (see *Nature* **406**, 555; 2000), as revealed in a recent survey of Marie Curie fellows. Funded by the European Commission (EC), these fellowships are aimed at supporting the training and mobility of researchers in Europe.

That these postdoc positions can be fraught with difficulties was brought into sharp focus by the survey of 200 Marie Curie Fellowship Association members (about 10% of its membership of past and present fellows). Set up by the EC to track and support fellows, the MCFA is independent and is now actively lobbying on behalf of the postdocs.

Travel guides

Marie Curie fellowships cost the EC 168 million euros (US\$144.5) in 1999, the first



Career directions: the destinations for the Marie Curie postdocs in 1999.

year of the EC's current research programme. Almost 1,500 individual researchers applied, and around 600 were recommended for funding. Most postdocs chose to go to the United Kingdom, and then France and Germany (see chart). The median age for fellows is 28, and many (40%) work in the life sciences.

Although most fellows are satisfied with

the scientific training, experience and confidence gained by spending two years abroad, there can be a high cost attached. The mobility problems are just the tip of the iceberg. Fellows from Spain and Germany say their home countries are reluctant or unwilling to recognize titles or qualifications gained abroad. Many postdocs taking positions outside their home country struggle to

Finding a focus for European postdocs

Although US postdocs are already well organized, with associations starting up across the country, in Europe there has been something of a vacuum. The vast majority of postdocs seek permanent positions but only 5–10% will succeed; leaving most in a series of short-term contracts, little autonomy and no career development. European postdocs are getting a raw deal and, as much casual labour is employed in the science sector, this is likely to get worse.

But now it seems that there could be some good news on the horizon for European postdocs. The Marie Curie Fellowship Association (MCFA) — which has

groups in most European countries — has a bird's-eye view of both European postdocs and postdocs in different countries, and is starting to speak up on behalf of all postdocs. In November, the European Science Foundation (ESF) will hold a joint postdoc conference with the MCFA — picking up on an earlier ESF survey of the situation of European postdocs.

Tony Mayer, head of the secretary-general's office at the ESF, says that the organization has found it difficult to talk to a mass of young European researchers. "In most countries there is very little in the way of organization," he says. But the MCFA

represents an organized group of some of the more mobile European postdocs, he notes.

The ESF wants feedback from young researchers on issues for the future, such as the sixth Framework programme, and more on the plans for the European Research Area. "Inevitably the people involved in science policy are senior; it is partly time [that stops postdocs from being involved], and that we never tend to ask them," says Mayer. There will also be a web debate to allow postdocs to use the ESF as a sounding board over the European Research Area. **N.L.**
ESF ♦ <http://www.esf.org>



Trouble at home

Coming back to Spain after holding a postdoc position abroad, José Pablo Zamorano (above) finds himself at a career crossroads. Prospects for permanent employment in his native country appear bleak, and, at 35, he is approaching the age limit for a postdoc position in other EU countries.

The Spanish government has tried to minimize the postdoc 'brain drain' since 1992, when it established a scheme to find work for scientists who received postdoc training abroad. It hoped that giving them three-year 're-entry' positions would staunch the flow of Spanish scientists to other countries. But after participating in such a programme, Zamorano, like many other young Spanish scientists who sought training abroad, finds himself looking for work in other countries again — permanent public-sector jobs are still scarce in Spain.

New policies do not seem to be addressing the need for more permanent posts. Under the new National Plan for Research, Development and Technological Innovation, the government has pledged to create 2,000 new jobs in the public sector for newly qualified and senior postdocs during 2000–03. But contracts range from five years for newly qualified scientists, to ten years for more senior investigators. Jordi Petriz, a postdoc who had to leave the Cancer Research Institute in Barcelona when his temporary contract ended, received wide coverage in the Spanish media in July because the end of his contract coincided with the publication of his article in *Nature Medicine*. Petriz, an expert in the functional aspects of stem cells, believes the situation for Spanish postdocs is worsening. "There are really bright people abroad who cannot return to Spain," he says.

Spaniard Valentin Fuster, director of the Department of Cardiology at the Mount Sinai School of Medicine, New York, says that the situation in Spain is the result of "a complete absence" of long-term planning of biomedical research by the government. But he hopes that the creation of the new Ministry of Science and Technology will help to improve the scientific policy in the country.

Less bureaucracy would also help, says Zamorano. He noticed that in England, where he did a postdoc at the University of Nottingham, local research directors had far more control over how many researchers they hired. Research directors also have more autonomy in choosing their own postdocs. "Here, the contract is not decided by the group but by the state," Zamorano says. But there is one promising sign. On 8 September Rolf Tarrach took over as the new president of Spain's main research organization, the Higher Council of Scientific Research. Tarrach regards recruiting postdocs who have received training abroad as an urgent priority. ■

Xavier Bosch is *Nature's* contributing correspondent for Spain.

Many European universities still have a closed-shop attitude favouring the local candidates.

pounds, the salary is not very good and not enough to live on in the UK," he says. Fellows with dependent families in the United Kingdom are particularly hard up.

Many fellows say that dealing with the administrative issues in a foreign country was simply too great a problem and took a lot of time. "I lost too much time with administrative questions and social security problems," says one fellow who participated in the anonymous survey. "Europe is not yet ripe for such a mobility." And some fellows say host institutions, which manage the fellowship fund and pay the fellow's salary, are taking money that they are not entitled to have. Another anonymous MCFA fellow writes: "We need rigid, enforceable guidelines on what the host institution can and cannot do with the money."

Administration problems

Meinhard Ober, a researcher in the materials testing institute at the University of Stuttgart, and until recently the national coordinator of the German MCFA group, says: "Most of the problems are with the administration of the university and the foreign country. The administration problems come from the fact that the number of fellows is very small [in any one institution]." A recent MCFA conference in Germany proposed that all fellows in one country should be employed by a single company — to ensure standard terms and conditions.

Emilios Harlaftis, acting chairman of the Greek MCFA and an assistant professor at the Institute of Astronomy and Astrophysics in Athens, faced similar administrative problems when he took up his fellowship in the UK. He agrees that companies should be administering these fellowships. "Clear and practical directives to the national institutes and universities are missing," he says.

And the news is not necessarily good for fellows returning home. Many MCFA fellows complain they have lost all contacts with networks at home, face unemployment when they return, and are in a worse position than when they left because local candidates are favoured over those from abroad (see "Trouble at home", above).

Patrik Floreen, an EC scientific officer with expertise on Marie Curie fellowships, says mobility problems are not in the hands of the commission. "All these things relate to national law," he explains. Taxation and

► comply with national regulations requiring residency or working permits — usually because their institution or department had issued the wrong kind of contract.

Money can also be a problem. When researchers apply for a fellowship in another country, they have no idea what their actual monthly salary will be when they arrive. The EC offers a fixed fund for the fellowship, with a monthly rate for the salary, which varies between EC member states. But the fellow's salary will be affected by local taxation, social security and pension schemes, and these in turn depend on the fine detail of the contract

that the university or research centre draws up. Although model national contracts are available, there is no obligation to stick to them. In the Netherlands, for example, contracts vary from faculty to faculty.

In Britain, where a third of Marie Curie postdocs take up their fellowships, many fellows have suffered deep salary cuts because of the decline in the value of the euro. Juan-Antonio Gilabert, a Marie Curie fellow in molecular biology at the University of Oxford, says that Oxford is one of the most expensive places to live in Britain. "After transformation of your euro salary into

social security are dealt with by the member states. "The community does not have the right to go and harmonize. We know very well about these difficulties, we have years of experience but the problems are not caused by community-level trouble," he says.

In January this year, Philippe Busquin, the commissioner for research, revealed his plans for a European scientific community dubbed the European Research Area. One of its priorities will be tackling these barriers to mobility for European researchers.

Campbell Warden, the chairperson of the European Association of Research Managers and Administrators, which campaigns to improve research management, says: "National authorities should take their responsibilities now and start working together to even out these barriers." Warden also mentions: "Many European universities

(and other research institutions) still have a very closed-shop attitude towards favouring the local candidates. This is a matter of serious concern and should be fought with all means."

What is remarkable is that despite the weight of difficulties faced by many Marie Curie fellows, the scheme has achieved notable successes. Harlaftis says the fellowship worked out very well in the end and was important to his career. "I got invaluable technical and scientific know-how in one of the major European observatories, and so made possible my return to Greece by funding a one-year return fellowship. This placed me in a very strong position. I am now competing for the director's position in my institute," he says. ■

Natasha Loder is a news reporter at Nature.

Marie Curie fellowships ► <http://www.cordis.lu/improving>
MCFA ► <http://www.mariecurie.org>

Postdocs reject academic research

With full-time positions in universities at a premium, postdocs are looking to government labs and industry for jobs, says Potter Wickware.

"Many are called but few are chosen" might be the epigraph of many a US postdoc's search for a tenure-track position in academic research. Although a relatively small number are successful in getting such jobs in universities, too many find themselves consigned to

an academic underclass, with low pay, inadequate benefits and no clear scientific career prospects.

The numbers paint a bleak picture. The 1998 National Research Council report *Trends in the Early Careers of Life Scientists* puts the success rate for recent PhDs at

22.9% in 1997, down from 26.6% in 1993. And the growing supply of new PhDs continues to exceed the number of available positions. A recent survey in the *Chronicle of Higher Education* estimates the number of natural sciences PhDs in 1996–97 at nearly 9,300. If every one of them claimed an existing (but filled) full-time job, they would represent more than 9% of the full-time natural sciences academic faculty.

Alternative routes

With such slim pickings, astute young scientists are looking elsewhere to launch their careers. National laboratories, technical positions and biotechnology companies all hold promise.

Steven Sloop, one of about 200 postdocs at the Lawrence Berkeley National Laboratory in California, says that government labs provide better postdoc prospects than the academic sector. Long-established collaborations between industry and the national labs give postdocs there a good view of the government-private industry interaction. They also provide opportunities to learn about companies and meet their scientific representatives, says Sloop, a chemist who researches electrolytes for lithium batteries.

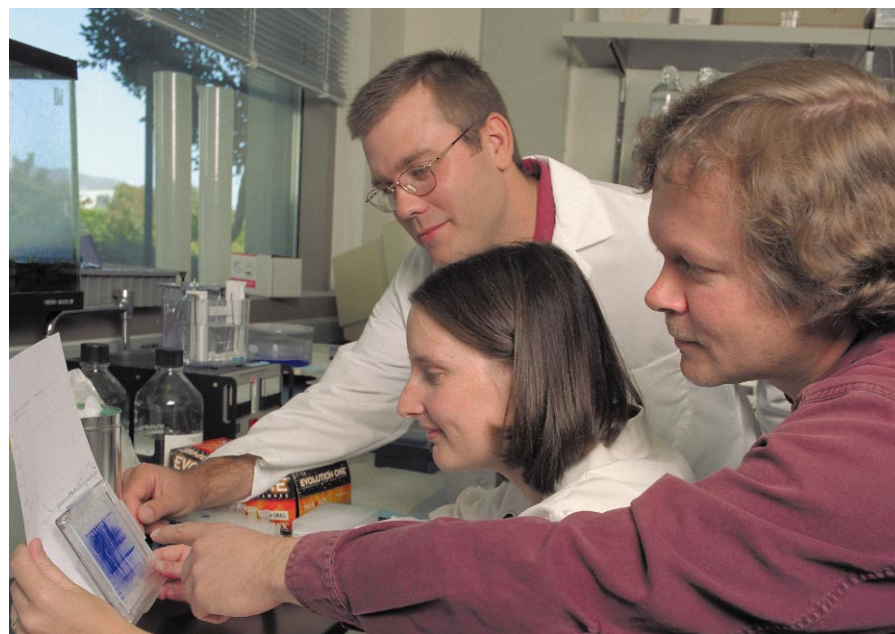
Sloop, who has been active in the lab's postdoc society, says that information sharing can break the isolation that all too easily envelops the researcher sequestered in a lab. Organizations also provide useful workshops in interviewing. And when different disciplines meet — materials scientists and molecular biologists, for example — new combinations of ideas can lead in unexpected and fruitful scientific directions.

Sloop plans to take his career in such a direction when his fellowship ends. He is making plans to move to Curitiba in southern Brazil. As the city has some of the most advanced car manufacturing plants in the world, Sloop thinks opportunities exist for hi-tech consultancy work.

The move is not without risk, but as much of the scientific talent in South America is flowing northwards, it is a calculated one. Sloop thinks his chances are good compared with those who stay at home and aim for more conventional goals. "With 200 or more applications from qualified people for a single faculty opening in a research university, any reasonable person develops a contingency plan," Sloop says. "Mine is Brazil."

Sloop's own career odyssey has placed him in a good position to give advice to new postdocs. He advises the prospective postdoc to be scrupulous about researching both the lab and the principal investigator beforehand. "This'll tell you if there is a history of problems, and even if the request raises eyebrows, how they respond will be informative about the atmosphere of the lab," he says.

Postdocs should also ask for periodic reviews of their own work, he believes. This is



Good prospects: many postdocs are turning to the biotech industry for employment.

common in industry but less so in most academic labs. Regular feedback is important so that the postdoc does not get overlooked or taken for granted in the life and work of the lab. "Be willing to ask questions; don't automatically go to the biggest name, but go instead to the best person. Similar philosophies are important. I ask the principal investigator why they became a scientist," he says.

Avoiding pitfalls

Some new PhDs might think twice about doing a postdoc at all. For the candidate set on an academic job, it is probably indispensable, but for those who want a job in industry or in a government lab, it may be superfluous. One good reason to do a postdoc is to tie up the loose ends from a PhD. Another reason might be to do a two-year post-PhD stint in a government or industrial setting to get a close look at government labs and companies. "One can get a lot of visibility in a postdoc fellowship," Sloop continues. "But don't stay too long."

Nancy McNamara prefers her current research assistant position at the University of California at San Francisco (UCSF) to an earlier postdoc post she held there. It seems "more like a real job", she says, because it includes benefits such as sick leave and vacation pay. McNamara, a member of the postdoctoral association, cites past abuses of postdocs at the UCSF, including instances of foreign doctors working as postdoc volunteers for no pay, as reasons to be wary. But she believes conditions for postdocs are improving slowly but perceptibly.

To avoid exploitation, potential postdocs should exercise due diligence before taking a job. "There are some awful principal investigators out there," McNamara says. "Talk to people in the lab, and people who have left the lab. How did they do afterwards? How do they feel about the experience? Are they going on to another postdoc somewhere or are they getting real jobs in industry or universities? Use the web, and network as much as possible."



Sloop: postdocs provide visibility.

Organization pays



After a disagreement with her principal investigator escalated into a summary dismissal from her post, and physical removal by university police from her lab, Joni Seeling (left) tried to get redress from the University of Utah Health Sciences Center. Seeling, who works at the university's oncology department, held on to her job after she retained an attorney. During her fight, she became increasingly disillusioned with the university's lack of mediation or appeals machinery, and what she saw as its inadequate levels of response to postdoc concerns.

Discovering that she was not alone among the university's 400 postdocs in having problems with her principal investigator, she set up a postdoctoral association in 1997. One of the group's main goals is to try to get written contracts for postdocs, which would spell out salary rates and a schedule for rises, length of employment, and benefits and retirement contributions. "Many postdocs are in their mid-thirties, even in their

forties, and have never received retirement benefits," she says.

She would also like to see a statement of what work postdocs can take with them when they leave to start their own lab or go to their next job, and how authorship priority is assigned in papers that emerge from the lab.

Seeling anticipated that her situation would improve after one of her papers appeared in *Science*, and when she received a National Institutes of Health investigator-initiated R01 grant with a high reviewer-approval rating. But little has changed. An agreed-upon salary rise has not yet appeared, and relations with university administration and faculty have remained a little strained.

Once cautious about organizing postdocs and pressing the university to change because of possible ill effects on her career prospects, she says: "I'm now at the point of saying 'if it hurts me, so be it'." **P.W.**

John Kerr, Sloop's principal investigator at Lawrence Berkeley, observes that even today many new PhDs, apparently unaware of the demographic numbers, come in with an expectation of a tenure-track post. He partly blames faculty, who fail to inform students about today's academic track realities. It is an error that is perhaps understandable because many of today's senior faculty were employed during a time of plenty. They may not be as aware of the current reality and are probably poor sources of job-market information.

Kerr also views with disfavour the exploitation suffered by some graduate students and postdocs, particularly in life sciences. The onus is partly on the student to understand that he or she is being exploited and to do something about it. But at the same time, "you have young people in vulnerable situations and university faculty, administration, and particularly the funding agencies, must take their responsibility to them seriously."

Biotechs beckon

Biotech companies are an emerging source of higher-paying, more stable sources of postdocs than universities. With a good publication record and deep financial resources, Genentech in San Francisco is just one of many biotech companies targeted by postdocs.

Abraham de Vos, who runs a protein crystallography lab there and also directs the company's postdoc programme, says that in the early 1990s Genentech established a postdoc programme based on the academic model. It now has 62 postdocs augmenting a staff of 100 research scientists. For the most

part, the research done is indistinguishable from work that would be carried out in an academic lab, although postdocs are kept away from some of the highly proprietary projects, and publication can sometimes be delayed until after a patent application.

First-year postdocs at Genentech are paid US\$44,000, almost twice as much as some postdocs get at universities or the National Institutes of Health. With the exception of stock options, benefits are the same as for other employees: health insurance, a retirement plan, employee stock purchase and maternity leave.

In evaluating a prospective postdoc, de Vos says, "I look for successful research in the academic lab that they came from and a feeling that they have done a lot of the work themselves, that they can take intellectual ownership. You want to see the seeds of independence in their previous graduate school experience."

Potter Wickware is a science writer in San Francisco.

Web links

US National Research Council Report: "Trends in the Early Careers of Life Scientists" ♦ <http://www.nap.edu/readingroom/books/trends/>

Joni Seeling's Utah Postdoc Association

♦ <http://www.cc.utah.edu/~jmj2/who.html>

Am. Assn. of University Professors Report: "The Status of Non-Tenure-Track Faculty" ♦ <http://www.aaup.org/rbnonten.htm>

UC Berkeley Graduate Teaching Assistants' Strike Committee ♦ <http://www.uaw.org/publications/releases/1998/calif.html>

UC Berkeley Postdoc Association

♦ <http://postdoc.berkeley.edu/bpa/infores/index.htm>

US National Academy of Sciences report

♦ <http://www.nap.edu/catalog/9831.html>